INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2023IPPAXXXX

Thèse de doctorat

TELECOM Paris

IP PARIS

Contributions to stochastic analysis for non-diffusive structures

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 18 décembre 2023, par

CHRISTOPHE VUONG

Composition du Jury :

Laure Coutin Professeure, Université de Toulouse	Présidente
Nicolas Privault Professor, Nanyang Technological University	Rapporteur
Raphaël Lachièze-Rey Maître de conférences HDR, Université Paris Cité	Rapporteur
Hélène Halconruy Maîtresse de conférences, Télécom SudParis	Examinateur
Laurent Decreusefond Professeur, Télécom Paris	Directeur de thèse
Marc Glisse Chargé de Recherche, INRIA Saclay - Île de France	Co-directeur de thèse

Contents

Remerciements

Je remercie mon directeur de thèse Laurent Decreusefond.

Introduction

Nous présentons les notions et outils nécessaires pour comprendre l'ensemble du manuscrit.

Diffusion

Les *diffusions* sont des fonctions aléatoires, qui sont très utilisées en physique, chimie, biologie, statistique et en finance. Leur nature même en fait un outil de modélisation formidable : elle permet de capter des dynamiques instantanées entachées d'incertitude. Au-delà de leur intérêt descriptif, elles se prêtent aux utilisations quantitatives. D'un point de vue probabiliste, ce sont des processus solutions d'équations différentielles stochastiques d'un certain type. Hormis les processus gaussiens, il n'est pas possible de les décrire en spécifiant les marginales fini-dimensionnelles. Les diffusions connues peuvent être décrites selon leur dynamique, par exemple l'équation de Langevin:

$$\mathrm{d}Y(t) = \sigma \ \mathrm{d}B(t) - bY(t) \ \mathrm{d}t.$$

où σ et b sont des paramètres réels de l'équation tandis que B est un mouvement Brownien. Pour simplifier nos propos dans cette introduction, on considère que le processus est à valeurs dans $\mathbb{R}^d$. Il est connu que le mouvement Brownien tout comme les diffusions sont des processus de Markov. Un processus de Markov Y est caractérisé par son opérateur de semi-groupe P. En le laissant agir sur une classe suffisamment riche de fonctions $f : \mathbb{R}^d \to \mathbb{R}$ boréliennes bornées :

$$(P_t \, f)(x) := \mathbb{E}\left[f(Y(t)) \,|\, Y(0) = x\right].$$

En général, on considère un espace de Banach séparable, muni de la norme $\|\cdot\|_\infty$, de fonctions continues tendant vers 0 à l'infini sur lesquelles le semi-groupe agit, dénoté $\mathcal{C}_0 = \mathcal{C}_0(\mathbb{R}^d, \mathbb{R})$. Nous rappelons le théorème suivant.

Theorem 0.0.1 *$(P_t, t \geq 0)$ est un semi-groupe conjointement continu de contractions sur $\mathcal{C}_0$, i.e.,*

1. $\|P_t \, f\|_\infty \leq \|f\|_\infty, \ \forall t, f$
2. $P_{t+s} = P_t \circ P_s$
3. $P_0 = \mathrm{Id}$,
4. *$(t, f) \mapsto P_t \, f$ est continu de $\mathbb{R}_+ \times \mathcal{C}_0 \mapsto \mathcal{C}_0$.*

Ainsi, on définit le domaine du générateur comme le sous-espace vectoriel

$$\mathrm{Dom}\, L := \{f \in \mathcal{C}_0 : \lim_{t \searrow 0} \frac{1}{t}(P_t \, f - f) \text{ existe}\},$$

et pour $f \in \operatorname{Dom} L$, on définit $L\,f$ par la valeur de cette limite. L'opérateur linéaire L est appelé *générateurinfinitésimal*. Le théorème suivant résulte du théorème de Hille-Yosida (théorème 2.6 dans Ethier and Kurtz (1986)).

Theorem 0.0.2 *1.* $\operatorname{Dom} L$ *est dense dans* $(\mathcal{C}_0, \|\cdot\|_\infty)$*;*

2. P_t *laisse stable* $\operatorname{Dom} L$ *:* $P_t(\operatorname{Dom} L) \subset \operatorname{Dom} L$*;*

3. $\forall f \in \operatorname{Dom} L,\ \frac{\mathrm{d}}{\mathrm{d}t} P_t\, f = P_t\, L\, f = L P_t f.$

Puisque P_t est solution de l'équation différentielle $\frac{\mathrm{d}}{\mathrm{d}t} P_t\, f = L P_t f$, on écrit formellement:

$$P_t = \exp(tL).$$

Par l'inégalité de Jensen, pour toute fonction convexe $\phi : \mathbb{R} \to \mathbb{R}$, et tout $t \geq 0$, et toute fonction $f \in \mathcal{C}_0$,

$$P_t(\phi(f)) \geq \phi(P_t\, f). \tag{0.0.1}$$

Par différentiation, en passant à la limite pour $t \to 0$, on a que:

$$L\phi(f) \geq \phi'(f) L f. \tag{0.0.2}$$

Pour de nombreuses applications sur les processus de Markov, il est commode que (0.0.2) soit une égalité pour des fonctions ϕ suffisamment régulières, c'est-à-dire que:

$$L\phi(f) = \phi'(f) L f + \text{erreur}. \tag{0.0.3}$$

Pour écrire cette erreur dans le cadre d'une diffusion, on se réfère à la formule de transport suivante :

$$\begin{cases} \operatorname{Var} f(C) = \operatorname{Var}(C) f'^2(C) \\ \text{biais } f(C) = (\text{biais } C) f'(C) + \frac{1}{2} \operatorname{Var}(C) f''(C). \end{cases} \tag{0.0.4}$$

où C est une quantité aléatoire. L'intuition de Bouleau consiste à identifier le biais et l'erreur quadratique dans cette formule avec les opérateurs L et Γ défini ci-dessous, qui forment une structure de Dirichlet. L'*opérateur carré du champ* est défini par

$$\Gamma(f, g) := \frac{1}{2} \left(L(fg) - gLf - fLg \right)$$

pour $f, g \in \mathcal{A}$ une algèbre dans $\operatorname{Dom} L$ telle que le produit de fonctions fg est dans le domaine de L. Un générateur est dit *diffusif* si:

$$L\phi(f) = \phi'(f) L f + \phi''(f) \Gamma(f, f). \tag{0.0.5}$$

Le monographe Bakry et al. (2013) est dédié à l'étude des semi-groupes associés parmi lesquels le semi-groupe d'Ornstein-Uhlenbeck, le semi-groupe de Laguerre et le semi-groupe de Jacobi. D'importantes conséquences en découlent comme l'existence d'une inégalité de Poincaré et des théorèmes limites. En particulier pour ces derniers, les résultats sont obtenus en considérant la structure en chaos de l'espace de probabilité d'intérêt. L'approche rejoint celle du calcul de Malliavin.

Liens avec le calcul de Malliavin

Le calcul de Malliavin est un outil puissant qui permet d'étudier les propriétés des processus stochastiques, tels que les processus de Markov, les processus de Lévy et les processus de Wiener (voir Nualart (2006)). Il est basé sur la notion de dérivée stochastique, qui est une généralisation de la dérivée ordinaire aux processus stochastiques. Une structure de Malliavin est propre à l'espace de probabilité dans lequel est défini lesdits processus stochastiques. Nous nous référons au chapitre 1 de Nourdin and Peccati (2012) pour un exemple de structure de Malliavin dans le cas unidimensionnel. De manière générale, la construction d'une telle structure passe par la définition d'un gradient D, d'une divergence δ et de l'opérateur de nombre $\bar{L} = \delta D$ sur leurs domaines respectifs. Les opérateurs D et δ sont liés par la relation de dualité:

Introduire l'espace de Malliavin $\mathfrak{K}$

$$\mathbb{E}[\langle DF, \eta \rangle_{\mathfrak{K}}] = \mathbb{E}[F \delta U];$$

Par l'approche de semi-groupe,
Grâce à la décomposition en chaos, on peut définir la notion d'intégrale stochastique. Sur l'espace de Wiener, sa définition selon le calcul de Malliavin ne nécessite pas d'hypothèse sur l'adaptabilité contrairement au calcul d'Itô. Pour notre type d'application, une conséquence importante de la décomposition en chaos est la formule de multiplication qui se formule dans le cas gaussien sous cette forme.

Multiplicati formulae

A partir de ce résultat, grâce à la méthode de Malliavin-Stein (voir la monographie Nourdin and Peccati (2012)), on obtient un théorème limite quantitatif en terme de distance de totale variation:

$$d_{TV}(F, \mathcal{N}(0,1)) \leq C\sqrt{\mathbb{E}[F^4] - 3}.$$

Structures sans propriété de diffusion

Pourtant, les structures sans propriété de diffusion sont légions. Sans cette propriété de diffusion, de nombreux résultats connus dans le chaos de Wiener comme le théorème de quatrième moment (Azmoodeh et al., 2014) ne peuvent pas être étendus. Le point critique dans les preuves est l'utilisation d'une règle de dérivation (en chaîne).
Récemment, il a été question de quand même étendre des résultats à d'autres chaos, qui ne bénéficie pas de la diffusion du générateur.
C'est là qu'intervient Malliavin.

> Faire un parallèle Brownien, Poisson.

Plus que la nature discrète, c'est l'absence de propriété de diffusion sur l'espace qui est un frein pour dérouler le calcul de Malliavin-Dirichlet.
Les fonctionelles de processus de Poisson rentrent dans ce cadre étant donné que un processus de Poisson (respectivement une mesure de Poisson) peut être vu comme un élément de l'espace des configurations sur $\mathbb{R}^+$ (respectivement $\mathbb{R}^d$ avec $d \geq 1$).
Cette thèse se compose de deux parties distinctes.

Calcul de Malliavin pour les variables aléatoires conditionnellement indépendantes

La première partie de ce travail porte sur l'implémentation de la méthode Malliavin-Stein pour des variables aléatoires conditionnellement indépendantes avec en vue une application à la

quantification de théorèmes limites concernant les hypergraphes aléatoires.
Nous nous intéressons aux fonctionnelles F de séquence de variables aléatoires $X = (X_a)_{a\in A}$ indépendantes conditionnellement à une variables aléatoire Z. Le gradient discret dans notre structure de Malliavin est défini pour $a \in A$ par:

$$D_a F := F - \mathbb{E}\left[F | (X_b)_{b\in A\setminus\{a\}}, Z\right].$$

Il constitue avec l'opérateur de divergence δ et l'opérateur d'Ornstein-Uhlenbeck L la boîte à outils fournie par la structure de Malliavin. Comme dans les travaux de Decreusefond and Halconruy (2019), celle-ci découle de la *dynamique de Glauber* associée à l'opérateur Laplacien de la structure de Malliavin discrète se rapporte à un échantillonnage de Gibbs étalé dans le temps.
Nous pouvons effectivement décrire la loi jointe d'une séquence de variables aléatoires indépendantes, comme la loi liée à la mesure invariante d'un tel processus. Le *semi-groupe* associé à la dynamique permet de déduire une formule de covariance pour montrer des inégalités de concentration. La modification de la dynamique comme présentée dans la première partie du manuscrit permet d'avoir une version des résultats obtenus pour les séquences de variables aléatoires conditionnellement indépendantes.
Du fait de cette caractérisation supplémentaire de la loi jointe, on obtient des outils supplémentaires qui complètent la boîte à outils, venant de la structure de Dirichlet sous-jacente, à savoir l'opérateur carré du champ et les formes de Dirichlet. Étant donné la relation entre l'un et l'autre, les résultats seront formulés en terme de l'opérateur carré du champ:

$$\Gamma(F,G) := \frac{1}{2}\left\{\mathsf{L}(FG) - F\mathsf{L}G - G\mathsf{L}F\right\}$$

La *formule d'intégration par parties* qui en découle est primordiale pour dérouler la méthode de Stein :

$$\mathbb{E}[\Gamma(F,G)] = -\mathbb{E}[F\mathsf{L}G]. \tag{0.0.6}$$

Nous obtenons des théorèmes limites quantitatifs pour des U-statistics de variables conditionnellement indépendantes. Des théorèmes de quatrième moment qui prennent appui sur des articles fondateurs de calcul de Malliavin dans une *structure de Dirichlet* (Azmoodeh et al., 2014) peuvent être déduits permettant de fournir des vitesses de convergence d'approximation normale des statistiques d'hypergraphes aléatoires. Nous l'appliquons pour la preuve de la normalité asymptotique du comptage de motifs dans un hypergraphe aléatoire $\mathbb{T}_n$ où les triangles sont tirés indépendamment avec une certaine probabilité conditionnellement à la présence des arêtes dans un graphe Erdös-Rényi $\mathbb{G}(n, p_n)$ avec $p_n < 1$. C'est un résultat nouveau dans le sens où la structure conditionnellement indépendantes était un frein pour la déduction de vitesse de convergence. Un théorème central limite conditionnel pouvait être déduit avec les méthodes usuelles mais pas plus. Avec l'application au comptage de motifs pour un hypergraphe, la structure avec une couche d'arêtes indépendantes, et de une couche de triangles conditionnellement indépendantes permet de déduire un vrai théorème central limite pour l'estimateur, qui est nouveau, car l'hypergraphe aléatoire donné est bien différent de l'hypergraphe aléatoire Erdös-Rényi $\mathbb{H}_n = \mathbb{G}^{(3)}(n, p'_n)$ (Lovász, 2012, section 23.3). Si on tire les triangles, avec probabilité 1, et que $p_n = 1/2$ et $p'_n = 1/8$, les deux modèles ont la même densité d'arêtes 1/8, mais $\mathbb{H}_n$ est quasi-aléatoire alors que $\mathbb{T}_n$ ne l'est pas. Il a une petite intersection avec chaque 3-uniforme hypergraphe quasi-aléatoire par la proposition 23.10 (Lovász, 2012). Les deux modèles se ressemblent néanmoins par leur homogénéité, car finalement issu du même modèle de graphe Erdös-Rényi étendu différemment. Rien qu'avec une famille de graphe aléatoire, on peut créer plusieurs sous-familles d'hypergraphes aléatoires, ce qui donne une latitude de modélisation et de résultats de convergence en loi pour des estimateurs sur ces modèles.

Inversibilité dans le contexte de fonctionnelles de mesure de Poisson

La seconde partie porte sur l'extension de la notion d'inversibilité au

Manuscrit

Le manuscrit est organisé comme suit.
Pour la suite du manuscrit, j'ai pris le parti de mettre en valeur les résultats nouveaux avec des boîtes colorées, en espérant une lecture plus aisée.

Part I

Stein-Malliavin-Dirichlet method and applications to statistics in hypergraph theory

Chapter 1

Introduction

1.1 Background

The fundamental example of the type of result we address in the first part is the Berry-Esseen bound for the central limit theorem. Let $X_1, \ldots, X_n$ be i.i.d. random variables with $\mathbb{E}|X_1|^3 < \infty$, $\mathbb{E}[X_1] = 0$ and $\mathrm{Var}(X_1) = 1$, if Φ denotes the cumulative distribution function of a standard normal distribution and $W_n = \sum_{i=1}^n \frac{X_i}{\sqrt{n}}$, then there exists $C > 0$ such that:

$$|\mathbb{P}(W_n \leq x) - \Phi(x)| \leq \frac{C\,\mathbb{E}|X_1|^3}{\sqrt{n}}.$$

That quantifies the error in the classical central limit theorem. More generally, a central theme of probability theory is proving distributional limit theorems. For the purpose of approximation it is of interest to estimate the rate of convergence in such results. The common methods employed are method of moments, Fourier analysis and martingale theory, but in some important cases they lead to sub-optimal bound on the rate.

Stein's method is a technique that can quantify the error in the approximation of one distribution by another in a variety of metrics. The Stein's method for normal approximation was invented in the groundbreaking paper (Stein et al., 1972) published in 1972 and extended for Poisson convergence (Chen, 1975) a few years later. Stein's method has proved powerful in particular for deriving explicit sharp bounds on distributional distances even when the underlying random element consists of structures with dependence (Arras et al., 2020). Moreover, it has been developed by a growing community to tackle an enlarging collection of approximation problems including beta, binomial, gamma, multinomial, variance-gamma, Wishart, and many more.

The common denominator of most of them is the existence of a generator whose invariant measure is the target distribution. The generator approach was a novel approach to Stein's method introduced in Barbour (1990) (see also Barbour (1988); Gotze (1991)). Since then, it has been a common scheme for approximation. That is one of the first building blocks of the Malliavin-Stein method in which the semigroup associated to the generator plays an important role.

The powerful interactions of Malliavin calculus of variations and Stein's method were highlighted by Nourdin and Peccati. The seminal paper Nourdin and Peccati (2009) used the combination of Malliavin calculus and Stein's approach to obtain a rather simple proof of the striking fourth moment theorem for normal approximation, established earlier in Nualart et al. (2005). Hence, that line of search has been popular as the fourth moment phenomenon emerges in many

contexts including free probability, compressed sensing, time series analysis, stochastic geometry and motif estimation in random graphs.
The latest developments in computational statistics are summarized in Anastasiou et al. (2023). We also point to a website that regroups the papers using Stein's method over the years:

`https://sites.google.com/site/steinsmethod/articles`.

In this manuscript, we deal with a branch of Stein's method combined with Malliavin calculus enriched by the structure given by Dirichlet forms, honing on the application to fourth moment limit theorems.

1.2 Probability distances

For the taxonomy of probability metrics and their history, the essential reference is the monograph (Rachev, 1991) that presents a unified and comprehensive approach to the theory of the more than seventy known probability metrics and their applications. Many of them are variants, particular cases or extensions of the Wasserstein and Lévy-Prokhorov metrics. We ponder on the first type of probability distances. For two probability measures P and Q on a measurable space $(\mathcal{X}, \mathcal{A}, \mathbb{P})$, the probability metrics we consider have the form:

$$d(P,Q) = \sup_{h\in\mathcal{H}} \left| \int h \, \mathrm{d}P - \int h \, \mathrm{d}Q \right|, \tag{1.2.1}$$

where $\mathcal{H}$ is a space of test functions $f : \mathcal{X} \to \mathbb{R}$. In the following, we assume $\mathcal{X} = \mathbb{R}^d$ and $\mathcal{A} = \mathcal{B}(\mathbb{R}^d)$.

Example 1.2.1. For the Kolmogorov distance, the space is:

$$\mathcal{H} = \{\mathbb{1}_{(-\infty,x)} : x \in \mathbb{R}\}.$$

The Kolmogorov metric, denoted d_K, is the maximum distance between cumulative distribution functions, so a sequence of distributions converging to a fixed distribution in this metric implies weak convergence, although the converse is not true since weak convergence only implies pointwise convergence of cumulative distribution function (c.d.f. for short) at continuity points of the target c.d.f..

Example 1.2.2. For the total variation metric denoted d_{TV}, the space of test functions is: $\mathcal{H} = \{\mathbb{1}_A : A \in \mathcal{A}\}$. We use it for approximation by discrete distributions, for example the Poisson distribution.

Example 1.2.3. The Wasserstein distances is defined as the minimum cost of transporting P to Q, where the cost is measured by the distance that each mass unit must be moved. Mathematically, the Wasserstein distance is defined as follows:

$$d_{W_p}(P,Q) = \inf_{\gamma\in\Gamma(P,Q)} \int_{M\times M} c(x,y)\,\gamma(\,\mathrm{d}x,\,\mathrm{d}y)$$

where $\Gamma(P, Q)$ is the set of all couplings between P and Q and c is a cost function. In the literature, the p-Wasserstein distances refers to the cases where the cost functions are power of Euclidean distances, i.e. $c(x, y) = \|x - y\|^p$. By the Kantorovich-Rubinstein duality formula, the first-order Wasserstein distance can be written in the form (1.2.1) where the underlying space of functions is a subspace of Lipschitz-continuous functions: $\mathcal{H} = \{f \in \mathrm{Lip}(\mathcal{X}, \mathbb{R}) : \mathrm{Lip}(f) = 1\}$. It is a common metric, also called Kantorovich-Rubinstein distance (or norm), occurring in many contexts as optimal transport, partial differential equations and even surprising ones as quantification of some variants of the uncertainty principle in quantum physics. In the following, we refer to it as the 1-Wasserstein distance with the notation d_W.

Proposition 1.2.4 *Let μ is absolutely continuous with respect to the Lebesgue measure λ. If the Radon-Nikodym derivative $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}$ is bounded by C, then for any random variable W,*

$$d_K(\mathcal{L}_W, \mu) \leq \sqrt{2C\, d_W(\mathcal{L}_W, \mu)}.$$

In the remainder, d_W is the main distance for our distributional approximations. The following theorem illustrates the interest in considering that distance between each element of a family of probability measures $(\mu_n)_{n\in\mathbb{N}}$ and a target probability measure μ.

Theorem 1.2.5 *Let $n \in \mathbb{N}$, the following two properties are equivalent:*

1. $d_W(\mu_n, \mu) \to 0$.
2. $\int_{\mathcal{X}} f \, \mathrm{d}\mu_n \to \int_{\mathcal{X}} f \, \mathrm{d}\mu$ *for all bounded and continuous functions from $\mathcal{X}$ to $\mathbb{R}$.*

In our case, μ_n is the law of functionals F_n of sequences of random variables, and μ is the Gaussian distribution.

1.3 Stein's method principle

The Stein's method relies on the characterization of the target distribution in the scheme of approximation. It consists in the construction of so-called Stein identities or equivalently in our case Stein operators which act on a space of functions $\mathcal{H}_*$ not to confuse with the starting space of test functions.

Definition 1.3.1. *The Stein operator* L is defined such that for a given random variable Y,

$$\mathbb{E}[Lf(Y)] = 0 \quad \forall f \in \mathcal{H}_* \iff Y \text{ has distribution } Q. \tag{1.3.1}$$

The key step in Stein's method is to transform the expression of the quantity $\left|\int h \, \mathrm{d}P - \int h \, \mathrm{d}Q\right|$ of which we take the supremum over $\mathcal{H}$ in (1.2.1) into a Stein identity of the form $\mathbb{E}[L\,\varphi_h(Y)]$ for $\varphi_h \in \mathcal{H}_*$.

For the following, we fix $\mathcal{H} = \{f \in \mathrm{Lip}(\mathcal{X}, \mathbb{R}) : \mathrm{Lip}(f) = 1\}$. We will explain one way to obtain a Stein operator via the generator approach, using the formalism of Decreusefond (2015) which uses the underlying Dirichlet structure to characterize both the initial space and the target space.

The Stein operator turns out to be a piece of a Dirichlet structure.

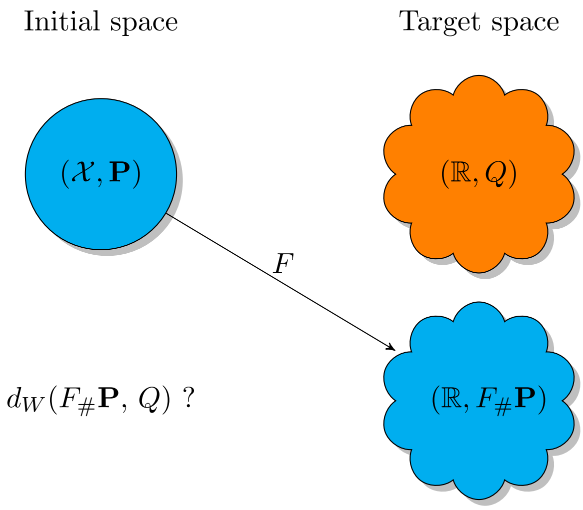


Figure 1.1: Comparison between $F_\# \mathbf{P}$ and Q.

Let E a metric space equipped with its σ-field $\mathcal{A}$, we recall some definitions.

Definition 1.3.2. A *Dirichlet structure* on $(E, \mathcal{A}, \mu)$ is a quadruple $(X^\circ, L^\circ, (P_t^\circ, t \geq 0), \mathcal{E}^\circ)$, where X° is a strong Feller process with values in E whose generator is L, and its semigroup is $(P_t, t \geq 0)$ for $f : E \to \mathbb{R}$ sufficiently regular.

In fact, we only need to define a Markov triple since the bilinear form $\mathcal{E}$ is not used here.

Definition 1.3.3 (Semigroup)**.** Let $(P_t)_{t \in \mathbb{R}^+}$ the family of operators defined on some set of real-valued measurable functions on $(E, \mathcal{A})$, and satisfying the following conditions:

1. For any $t \in \mathbb{R}^+$, P_t is a linear operator sending bounded measurable functions on $(E, \mathcal{A})$ to bounded measurable real functions.
2. $P_0 = \mathrm{Id}$ where Id is the identity operator (initial condition).
3. For every $(s, t) \in (\mathbb{R}^+)^2$, $P_{t+s} = P_t \circ P_s$.
4. For any $t \in \mathbb{R}^+$, P_t conserves the mass and preserves positivity (Markov property), i.e. $P_t \mathbb{1} = \mathbb{1}$. For any positive function f, $P_t f \geq 0$.

$(P_t)_{t \in \mathbb{R}^+}$ is a semigroup associated to a generator L.

Definition 1.3.4 (Invariant measure)**.** Let a family $(P_t)_{t \in \mathbb{R}^+}$ of operators defined on $(E, \mathcal{A})$ and satisfying all the properties above. A positive σ-finite measure ν_0 on $(E, \mathcal{A})$ is said to be *invariant* for $(P_t)_{t \in \mathbb{R}^+}$, if for every bounded function $f : E \to \mathbb{R}$, and $t \in \mathbb{R}^+$,

$$\int_E P_t f \, \mathrm{d}\nu_0 = \int_E f \, \mathrm{d}\nu_0. \tag{1.3.2}$$

In the following, we only consider semigroups that have invariant measure.

Remark 1.3.5. It must be noted that the knowledge of one of L°, P° or X° is equivalent to the knowledge of the other two.

We can use Dirichlet structures on both sides.

In the case of normal approximation, the associated Stein operator is for $f \in \mathcal{H}_*$:

$$Lf = f(x) - xf'(x). \tag{1.3.3}$$

There are several ways to obtain that elementary result. We recall an instructive one from Decreusefond (2015) which leverages the underlying Dirichlet structure associated to the normal distribution.

Definition 1.3.6. Let X° the Markov process defined as:

$$X^\circ(t,x) = \begin{cases} X^\circ(0,x) = x \\ \mathrm{d}X^\circ(t,x) = -X^\circ(t,x) + \sqrt{2}\,\mathrm{d}B(t), \end{cases} \tag{1.3.4}$$

where B is a standard Brownian motion. The solution to this stochastic differential equation is the Ornstein-Uhlenbeck process.

By the Itô formula, we have for $t \in \mathbb{R}^+$:

$$X^\circ(t,x) = xe^{-t} + \sqrt{2}\int_0^t e^{-(t-s)}\,\mathrm{d}B(s).$$

$X^\circ(\cdot,x)$ is then a Gaussian process with parameters:

$$m(t) = xe^{-t} \quad K(s,t) = e^{-|t-s|} - e^{-(t+s)}. \tag{1.3.5}$$

Hence, $X^\circ(t,x) \sim \mathcal{N}(xe^{-t}, 1-e^{-2t})$. We have that $X^\circ(t,x) \implies \mathcal{N}(0,1)$ when $t \to \infty$. The invariant distribution is the $\mu = \mathcal{N}(0,1)$.
Let $f \in \mathcal{S}$ the set of functions belonging to $\mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ such that f and all its derivatives have at most polynomial growth, and $x \in \mathbb{R}$. By the Mehler representation formula, the associated semigroup is defined as:

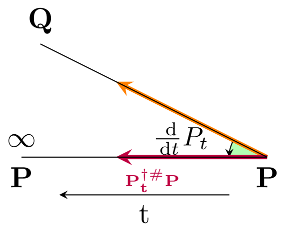
$$P_t f(x) = \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right)\mu(\,\mathrm{d}y). \tag{1.3.6}$$

We want to compute $\frac{d}{dt}P_t|_{t=0}$, for the Gaussian distribution μ as target. For any $f \in \mathcal{S}$,

$$\begin{aligned}\frac{d}{dt}P_t f(x) &= -xe^{-t}\int f'(e^{-t}x + \sqrt{1-e^{2t}}y)d\mu(y) + \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\int f'(e^{-t}x + \sqrt{1-e^{-2t}}y)y\,\mathrm{d}\mu(y) \\ &= -xe^{-t}\int f'(e^{-t}x + \sqrt{1-e^{2t}}y)\,\mathrm{d}\mu(y) + e^{-2t}\int f''(e^{-t}x + \sqrt{1-e^{-2t}}y)\,\mathrm{d}\mu(y) \text{ using lemma.}\end{aligned} \tag{1.3.7}$$

In particular, $\frac{d}{dt}P_t f(x)|_{t=0} = -xf'(x) + f''(x)$. We want to prove that $\mathrm{L}f = \frac{d}{dt}P_t f(x)|_{t=0}$. We would denote first $\mathrm{L}_t f(x) = \frac{d}{dt}P_t f(x)$

Which lemma?



Title for picture

We also write:

$$P_t f(x) - P_0 f(x) = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} P_s f(x)\ \mathrm{d}s = P_t f(x) - f(x). \tag{1.3.8}$$

$$\lim_{t\to\infty} \mathbb{E}[f(\mathcal{N}(e^{-t}x, 1-e^{-2t}))] - f(x) = \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}s} P_s f(x)\ \mathrm{d}s. \tag{1.3.9}$$

Hence, for any measure of probability ν,

$$\int f(x)\ \mathrm{d}\mu(x) - \int f(x)\ \mathrm{d}\nu(x) = \int \int_0^{+\infty} \mathrm{L}_s f(x)\ \mathrm{d}s\ \mathrm{d}\nu(x). \tag{1.3.10}$$

One wants to derive under the integral. Thus,

$$d(\mu,\nu) = \sup_{\varphi_h:\, h\in\mathcal{H}} \left| \int \mathrm{L}_t \varphi_h \right|. \tag{1.3.11}$$

The upper bound is achieved by functions solutions of Stein equations.
For $f \in \mathrm{Dom}(L^\dagger)$:

$$L^\dagger f(x) = x f(x) - f'(x). \tag{1.3.12}$$

Remind all derivatives semi-group and rewrite the space of test functions

$f^\dagger$ all order derivatives (see Nourdin and Peccati (2012, chapter 3 p.63-69)).
Especially,

$$d(F, \mathcal{N}(0,1)) \le \sup_{f^\dagger \in \mathcal{H}_*} |\mathbb{E}[L^\dagger f^\dagger F]|, \tag{1.3.13}$$

- $\mathcal{H}_{TV} = \{f^\dagger :\ \|(f^\dagger)'\|_\infty \le \sqrt{\pi/2}, \|f^{\dagger''}\|_\infty \le 2\}$;
- $\mathcal{H}_{Kol} = \{f^\dagger : \|f^\dagger\|_\infty \le \sqrt{2\pi}/4, \|(f^\dagger)'\|_\infty \le 1\}$

Mention Stein's equations based on Stein's operator

Takeaway message:
Reminder of Malliavin-Stein-Dirichlet method. All roads lead to Rome. (Decreusefond, 2015).
In the following section, we recall the principle of Malliavin-Stein's method, which is a ramification of Stein's method with several breakthroughs concerning normal approximation (see Nourdin and Peccati (2012)).
The first step of Stein's method relies on the characterization of the target distribution. The second step boils down to bound (1.3.13). There are various approaches to achieve that, using the structures of the random variables.

1.4 Malliavin-Stein-Dirichlet

Malliavin calculus is also known as the stochastic calculus of variations. At the very core of it, it considers a gradient on a measured space. The link between these, the differential geometry and the measure is made through the so-called integration by parts formula. When the measured space is the Wiener space, i.e. the set of continuous functions with the Brownian measure, the gradient generalizes the usual gradient on $\mathbb{R}^N$ and the integration by parts yields an extension of the Itô integral. The intersection with Stein's method originates from the seminal paper by Nourdin and Peccati, who were able to associate a quantitative bound to the remarkable fourth moment theorem on the Wiener space established by Nualart and Peccati (Nualart et al., 2005).

Fourth moment theorems are simplifications of results using the Method of Moments. The Method of Moments in probability theory is one of the oldest versatile tool used by probabilists to prove limit theorems in non-standard problems. The principle stems from the property that the moments of a random variable are determined by the distribution. Although the converse is not true, we have that the distribution of some random variable X is determined by its moments if X has finite moments and every random variable with the same moments as X has the same distribution. The standard version of the method of moments can be stated as follows (Chung, 1974, Theorem 4.5.5).

Theorem 1.4.1 *Let Z be a random variable with a distribution that is determined by its moments. If $X_1, X_2, \dots$ are random variables with finite moments such that $\mathbb{E}[X_k^n] \to \mathbb{E}[Z^k]$ as $n \to \infty$ for every integer $k \geq 1$, then $X_n \xrightarrow{d} Z$.*

It has been used for results of asymptotic normality of statistics on random graphs (see Janson et al. (2000, chapter 6)), and Poisson convergence as well. Overall, it is well adapted to combinatorial problems. The drawback is that it usually requires tedious calculations for the estimations. In some cases, the proofs can be reduced to the control of some moments and not all of them. This is the case of the fourth limit theorem.

Definition 1.4.2 (Cumulants)**.** Let F be a real-valued random variable such that $\mathbb{E}[|F|^m] < \infty$ for some integer $m \geq 1$, and write $\varphi_F(t) = \mathbb{E}[e^{itF}]$, $t \in \mathbb{R}$, for the characteristic function of F. Then, for $r = 1, \dots, m$, the *r-th cumulant* of F denoted by $\kappa_r(F)$, is given by

$$\kappa_r(F) = (-i)^r \frac{\mathrm{d}^r}{\mathrm{d}t^r} \log \varphi_F(t) \tag{1.4.1}$$

Remark 1.4.3. When $\mathbb{E}[F] = 0$, then the first four cumulants of F are the following: $\kappa_1(F) = \mathbb{E}[F] = 0$, $\kappa_2(F) = \mathbb{E}[F^2] = \mathrm{Var}(F), \kappa_3(F) = \mathbb{E}[F^3]$, and:

$$\kappa_4(F) = \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2.$$

The fourth moment theorem states a rate of convergence in terms of cumulants for a sequence of multiple stochastic Wiener-Itô integrals of fixed order $I_q(f)$ for $q \geq 2$.

Theorem 1.4.4 *Let $(F_n)_{n\geq 1} = (I_q(f_n))_{n\geq 1}$ be a sequence of random elements in a fixed Wiener chaos of order $q \geq 2$ such that $\mathbb{E}[F_n^2] = q!\|f_n\|^2 = 1$. Then, as n tends to infinity, the following assertions are equivalent.*

1. *$F_n \to \mathcal{N}(0,1)$ in distribution*

2. *$\kappa_4(F) \to 0 = \kappa_4(N)$*

where $N \sim \mathcal{N}(0,1)$.

A few years later, Nourdin and Peccati (Nourdin and Peccati, 2009) made a huge breakthrough by providing an error bound in terms of the fourth moment. The fourth moment phenomenon has since then emerged as a unifying principle governing the central limit theorems for various non-linear functionals of random fields. It has paved the way for a fruitful line of search, growing as far as obtaining the *universality phenomenon* according to which the asymptotic behavior of large random systems does not depend on the distribution of its components (see the comprehensive book Nourdin and Peccati (2012)). More precisely, one can use random variables in Wiener chaos instead of independent random variables in order to prove asymptotic normality of functionals of those provided some mild assumptions on the functionals. The intuition behind also led to a connection of the Malliavin-Stein's method and criterions of asymptotic normality in the literature such as DeJong's one (de Jong, 1989) which states a partial fourth moment theorem concerning multilinear forms, a particular class of functionals of independent random variables. The additional remainder in the convergence rate is expressed in terms of the maximal influence of the variables, roughly speaking the maximum over $i \in \mathbb{N}$ contribution of a random variable X_i to the overall configuration of the multilinear forms. Due to the discrete nature, the techniques used for the Wiener chaos do not apply in a straightforward way, although they can be adapted. It is the object of the remainder and of my contribution detailed in the next chapter.

1.5 The Markov triple approach

In this section, we introduce a general framework for studying the fourth moment phenomenon with chaos. The forthcoming approach was first introduced in Ledoux et al. (2012), and then further developed in Azmoodeh et al. (2014, 2016).
The main assumption is the property of *diffusion* of the generator L on the initial space.

Definition 1.5.1. A Markov operator L is said to be a *diffusive*, if for every function $\varphi : \mathbb{R} \to \mathbb{R}$, and for every smooth (enough) function $\varphi : \mathbb{R} \to \mathbb{R}$, and for every function $f \in \mathcal{H}$,

$$L\varphi(f) = \varphi'(f)Lf + \varphi''(f)\Gamma(f, f). \tag{1.5.1}$$

Equivalently, the associated carré du champ operator Γ is a derivation in the sense that: $\Gamma(\varphi(X), X) = \varphi'(X)\Gamma(X, X)$ for any $\varphi \in \mathcal{C}^\infty(\mathbb{R})$.

That is a consequence of the chain rule, which leads to:

$$\Gamma(\psi(f), \theta(g)) = \psi'(f)\theta'(g)\Gamma(f, g). \tag{1.5.2}$$

Definition 1.5.2. A fourth moment structure is a triple (E, μ, L) such that:

1. (E, μ) is a probability space;
2. L is a symmetric unbounded operator defined on some dense subset of $L^2(E, \mu)$ that we denote $\operatorname{Dom} L$, the domain of L°
3. The associated carré-du-champ operator Γ is a symmetric bilinear operator defined by:

$$\Gamma(F, G) := \frac{1}{2}\left\{L(FG) - FLG - GLF\right\}$$

4. The Markov operator L is diffusive.

Definition 1.5.3 (Azmoodeh et al. (2014) p.6)**.** An eigenfunction X of the generator $-L$ with eigenvalue λ_p is called a *chaos eigenfunction of order* p, if and only if:

$$X^2 \in \bigoplus_{k=0}^{2p} \ker(L + \lambda_k \mathrm{Id}). \tag{1.5.3}$$

That holds in the three most important diffusion structures, namely Wiener, Laguerre and Jacobi and in discrete setting like the Poisson space (Azmoodeh et al., 2014).

Chain rule for Stein's method

1.6 Fourth moment on the Poisson space

We introduce to an emblematic example of non-diffusive structure for which we can state the fourth moment using an analogous scheme of approximation. The Poisson space The following reminds some results in Döbler et al. (2018).
We recall the definition of the discrete gradient on the configuration space.

Definition 1.6.1. Let N be a Poisson process with intensity σ. Let $F : \mathfrak{N}_E \to \mathbb{R}$ be a measurable function such that $\mathbb{E}[F(N)^2] < \infty$. We define $\operatorname{Dom} D$ as the set of square integrable random variables such that

$$\mathbb{E}\left[\int_E |F(N \oplus x) - F(N)|^2 \, \mathrm{d}\sigma(x)\right]$$

This is a review of the existing fourth moment theorems that have been available in the literature. They have been obtained for most of them in the same manner, leveraging a chaos decomposition.

Rademacher setting

See Privault et al. (2008). The following chapter shows how using analogous techniques we can derive equivalent limit theorems as well as concentration results.

Completion of Malliavin recipe

Now we define the tools from Nicolas Bouleau, Guillaume Poly that unify Malliavin calculus of Nourdin, Peccati and Dirichlet computations. The definitions are from another insightful article about a similar approach (Ledoux et al., 2012). We recall that a Malliavin structure revolves around a gradient, its associated *Ornstein-Uhlenbeck operator*, and a *formula of integration by parts*.

The integration by parts given by the Dirichlet structure supersedes the integration by parts at the cornerstone of Malliavin structure. It is actually two ways of writing the integration by parts, at least in our case. The procedure is:

- Find an Ornstein-Uhlenbeck operator,
- gradient, difference operators that are the crux of the matter in computations.
- Integration by parts.

1.7 Contributions

The control by the fourth moment in a Wiener chaos and Poisson space to the quality of normal approximation in that setting is now well understood (Chen, 2021). The next natural random elements to deal with is the sequence of independent random variables. Lots of works about it including Decreusefond and Halconruy (2019) and Duerinckx (2021) about Malliavin calculus in that discrete setting.

There are limit theorems for subhypergraph counts in random hypergraphs (De Jong, 1996). However, most of them consider simply random hypergraphs as extended random graphs, hence considering hyperedges as independent random variables. In Austin (2008), the taxonomy of exchangeable random hypergraphs concern conditionally independent Bernoulli random variables. Hence, we have focused on the branch of Stein's method leveraging a Malliavin-Dirichlet structure for conditionally independent random variables. For an exhaustive overview of the Malliavin-Stein method, we refer the reader to the website:

`https://sites.google.com/site/malliavinstein/home`.

Add fourth moment theorem and the subhypergraph counts

In brief, our contributions are:

- A new Malliavin framework for conditionally independent random variables;
- An application to concentration inequalities;
- New bounds on the 1-Wasserstein bounds for U-statistics;
- An application to asymptotic normality of subhypergraph counts in random hypergraphs.

Chapter 2

Malliavin calculus for conditionally independent random variables

On any denumerable product of probability spaces, we extend the discrete Malliavin structure for conditionally independent random variables. As a consequence, we obtain the chaos decomposition for functionals of conditionally independent random variables. We also show how to derive some concentration results in that framework. The Malliavin-Stein method yields Berry-Esseen bounds for U-Statistics of such random variables. It leads to quantitative statements of conditional limit theorems: Lyapunov's central limit theorem, De Jong's limit theorem for multilinear forms. The latter is related to the fourth moment phenomenon. The final application consists of obtaining the rates of normal approximation for subhypergraph counts in random exchangeable hypergraphs including the Erdös-Rényi hypergraph model. The estimator of subhypergraph counts is an example of homogeneous sums for which we derive a new decomposition that extends the Hoeffding decomposition.

2.1 Motivation

It is only very recently that, concomitantly, the situation where the measured space is a product space, i.e. if we deal with independent random variables, has been addressed (see Duerinckx (2021); Decreusefond and Halconruy (2019); Dung (2018)). By order of complexity, the next situation which can be analyzed is that of conditionally independent random variables. This is a very common structure as de Finetti's theorem says that an infinite sequence of random variables is exchangeable if and only if these random variables are conditionally independent. This is the key theorem to develop a theory on random hypergraphs as in Austin (2008).
The first definitions of gradient (denoted by D) and divergence we introduce below for conditionally independent random variables, bear strong formal similarities with those of Decreusefond and Halconruy (2019). The difference lies into the computations which rely heavily on conditional distributions given the latent variable, which is here called Z. We can then follow the classical development of the Malliavin calculus apparatus: gradient, divergence, chaos, number operator and Ornstein-Uhlenbeck semi-group (denoted by P_t). We can even describe the dynamics of the Markov process whose infinitesimal generator is the number operator. At a formal level, the computations are almost identical to those of Decreusefond and Halconruy (2019) with expectations replaced by expectations given Z. That notion of gradient is itself the extension of gradient in the particular discrete setting when X is a sequence of Rademacher random variables Privault et al. (2008, section 10 proposition 10.1.).

Nevertheless, for more advanced applications, namely functionals identities like the covariance representation formula, we need to introduce a difference operator (see Definition 2.4.9) which appears more often that the gradient itself. It is in some sense a finer tool that the original gradient which is useful to define the Dirichlet structure (the Glauber process, the infinitesimal generator denoted by L, etc.) but no more. This is due to the fact that $D_a D_a = D_a$, which entails that L commutes with D and thus we have $DP_t = P_t D$ in place of the usual formula $DP_t = e^{-t} P_t D$ which is the core formula to derive all functional inequalities in the Gaussian and Poisson cases. The difference operator Δ allows to recover the crucial e^{-t} factor (see Proposition 2.5.1).

The prevailing application of Malliavin calculus is nowadays, the evaluations of convergence rates via the Stein's method (Nourdin and Peccati (2012); Decreusefond (2022) and references therein). The question is to assess a bound of the distance between a target distribution (more often the Gaussian distribution) and the law of a deterministic transformation of a probability measure, called the initial distribution.

One of the key difference between the Gaussian case and so-called discrete situations (Poisson, Rademacher, independent random variables) is the chain rule formula: it is only in the former framework that $D\psi(F) = \psi'(F)\, DF$. For the other contexts, we need to resort to an approximate chain rule Reinert et al. (2010). This is the role here of Lemma 2.6.7 and Lemma 2.7.1. Motivated by the applications to random graphs statistics, we focus here on normal approximations of U-statistics as in Barbour et al. (1989); Röllin (2022); Krokowski et al. (2017); Privault and Serafin (2018). In passing, we extend the notion of U-statistics by allowing the coefficients to depend on the latent variable instead of being only deterministic. Following the strategy of Azmoodeh et al. (2014), we establish a fourth moment theorem for such functionals.

The rest of the chapter is organized as follows. The section 2.2 lays the foundations of the Malliavin framework. We define the Malliavin operators and especially the gradient that is related to the so-called *Glauber dynamics*. Our focus respectively lies on the independent setting and the conditional independent setting (i.e. X is a sequence of independent random variables, respectively conditionally independent random variables). We follow the original approach of completing it with a Dirichlet structure (see for example Döbler et al. (2018, 2019)) that naturally arises without further assumption that gives us another formula of integration by parts with the *carré du champ operator*. We derive some functional identities in section 2.5, specifically conditional versions of Poincaré inequality and McDiarmid's inequality. The section 2.6 presents results of normal approximation. We will see that using the carré-du-champ operators instead of the norms of Malliavin gradient will allow us to bypass at once all combinatorial difficulties, leading to a partial fourth moment theorem for U-statistics under mild assumptions (see subsection 2.7) in the same vein of Azmoodeh et al. (2014) that shows the fourth moment theorem on the Wiener space.

2.2 Discrete Malliavin-Dirichlet structure

Let A be an at most denumerable set equipped with the counting measure, and define:

$$\ell^2(A) := \left\{ u \,:\, A \to \mathbb{R},\ \sum_{a\in A} |u_a|^2 < \infty \right\} \text{ and } \langle u, v\rangle_{\ell^2(A)} := \sum_{a\in A} u_a v_a.$$

Let $(\Omega, \mathcal{T}, \mathbb{P})$ be a probability space, E_0 be a Polish space and $((E_a, \Upsilon_a),\, a \in A)$ be a family of Polish spaces such that

$$\begin{aligned} E_A &= \prod_{a\in A} E_a \\ \Omega &= E_0 \times E_A \end{aligned} \tag{2.2.1}$$

The product probability space E_A is endowed with its Borel σ-algebra denoted $\Upsilon \subset \mathcal{T}$. Let Z a E_0-valued random variable. By Theorem 10.2.2 (Dudley, 2002), all the subsequent conditional distributions in the chapter admit regular versions. For any subset B of A, we denote the set $E_B := \prod_{b\in B} E_b$ and for $x \in E_A$, $x_B := (x_a, a \in B) \in E_B$ so that for $a \in B$, $x_a \in E_a$. We denote $x^B = (x_a, a \in A \setminus B)$. Let $X := (X_a)_{a\in A}$ be a sequence defined on $(\Omega, \mathcal{T}, \mathbb{P})$ of conditionally independent random variables given Z such that for all $a \in A$, X_a is an E_a-valued random variable, i.e.:

$$X_a \underset{Z}{\perp\!\!\!\perp} (X_b, b \in A \setminus \{a\}),$$

or, equivalently:

$$\mathbb{P}(X_a \in \cdot \,|\, \sigma((X_b, b \neq a), Z)) = \mathbb{P}(X_a \in \cdot \,|\, \sigma(Z)).$$

We denote by $\mathbf{P}$ the law of X and $\mathbf{P}^Z$ the law $\mathcal{L}(X|Z)$. See chapter 5 of Kallenberg (1997) for a thorough review of conditional independence, and Rao (2009) for some limit theorems for conditionally independent random variables. We use the notation $\mathbb{E}$ for the expectation of a random variable. By the disintegration theorem, for $a \in A$, the conditional probability distribution of X_a given $\sigma(X^{\{a\}}) \vee \sigma(Z)$ admits a regular version $\mathbf{P}_a$. For $p \geq 1$, let us denote $L^p(E_A \to \mathbb{R}, \mathbf{P})$ the set of p-th-integrable functions on E_A with respect to the measure $\mathbf{P}$. It is equipped with the norm $\|\cdot\|_{L^p(E_A\to\mathbb{R},\mathbf{P})}$ defined for f a measurable function on E_A by $\|f\|_{L^p(E_A\to\mathbb{R},\mathbf{P})} := \int |f(x)|^p \mathbf{P}(\,\mathrm{d}x)$. For the sake of notations, $L^p(E_A)$ stands for the space of p-integrable functionals

$$L^p(E_A) := \Big\{\omega \mapsto F(X(\omega)) : \omega \in \Omega, F \in L^p(E_A \to \mathbb{R}, \mathbf{P})\Big\}.$$

In this respect, $L^\infty(E_A)$ is the space of bounded functionals. We shall write F in place of $F(X)$ for the sake of conciseness. We closely follow the usual construction of Malliavin calculus on that space.

Definition 2.2.1. A functional F is said to be cylindrical if there exists a finite subset $I \subset A$ and a functional F_I in $L^2(E_I)$ such that $\mathbb{E}[|F_I|^2] < +\infty$ and $F = F_I \circ r_I$, where r_I is the restriction operator:

$$\begin{aligned} r_I \,:\, E_A &\longrightarrow E_I \\ (x_a, a \in A) &\longmapsto (x_a, a \in I). \end{aligned}$$

It is clear that the set of those functionals $\mathcal{S}$ is dense in $L^2(E_A)$. We set $L^2(A \times E_A)$ the Hilbert space of processes which are square-integrable with respect to the measure $\sum_{a\in A} \Delta^{\{a\}'} \otimes \mathbf{P}$, i.e.

$$L^2(A \times E_A) = \{U : \sum_{a\in A} \mathbb{E}\left[U_a(X)^2\right] < +\infty\},$$

equipped with the norm and inner product:

$$\|U\|_{L^2(A\times E_A)} := \sum_{a\in A} \mathbb{E}\left[U_a^2\right] \text{ and } \langle U, V\rangle_{L^2(A\times E_A)} := \sum_{a\in A} \mathbb{E}\left[U_a V_a\right].$$

Definition 2.2.2. The set of simple processes, denoted $\mathcal{S}_0(l^2(A))$ is the set of random variables defined on $A \times E_a$ of the form

$$U = \sum_{a \in A} U_a \mathbb{1}_a,$$

for $U_a \in \mathcal{S}$.

2.2.1 Malliavin operators

Definition 2.2.3 (Discrete gradient)**.** For $F \in \mathcal{S}$, DF is the simple process of $L^2(A \times E_A)$ defined for all $a \in A$ by:

$$D_a F := F - \mathbb{E}\left[F \,|\, X^{\{a\}}, Z\right].$$

In particular, $\mathcal{S} \subset \operatorname{Dom} D$. Define the σ-field $\sigma(X^{\{a\}}) \vee \sigma(Z)$ by $\mathcal{G}^a$, so that

$$D_a F = F - \mathbb{E}\left[F \,|\, \mathcal{G}^a\right]. \tag{2.2.2}$$

Recall that for $K \subset A$, $X_K = (X_a, a \in K)$ and $X^K = (X_a, a \in A \setminus K)$. We shall write $\mathcal{G}^K = \sigma(X^K) \vee \sigma(Z)$ and $\mathcal{G}_K = \sigma(X_K) \vee \sigma(Z)$ for K a subset of A.

Lemma 2.2.4 *Let $(a, b) \in A^2$, $a \neq b$, for $F \in \operatorname{Dom} D$,*

1. $D_a D_a F = D_a F$;
2. $D_a D_b F = D_b D_a F$;
3. $D_a \mathbb{E}\left[F \,|\, \mathcal{G}^b\right] = D_b \mathbb{E}\left[F \,|\, \mathcal{G}^a\right]$.

Proof of lemma 2.2.4. For $(a, b) \in A^2$, with $b \neq a$,

$$\begin{aligned}
D_a D_b F &= D_b F - \mathbb{E}\left[D_b F \,|\, \mathcal{G}^a\right] \\
&= F - \mathbb{E}\left[F \,|\, \mathcal{G}^b\right] - \mathbb{E}\left[F \,|\, \mathcal{G}^a\right] + \mathbb{E}\left[\mathbb{E}\left[F \,|\, \mathcal{G}^b\right] | \mathcal{G}^a\right] \\
D_b D_a F &= D_a F - \mathbb{E}\left[D_a F \,|\, \mathcal{G}^b\right] + \mathbb{E}\left[\mathbb{E}\left[F \,|\, \mathcal{G}^a\right] | \mathcal{G}^b\right] \\
&= F - \mathbb{E}\left[F \,|\, \mathcal{G}^a\right] - \mathbb{E}\left[F \,|\, \mathcal{G}^b\right] + \mathbb{E}\left[\mathbb{E}\left[F \,|\, \mathcal{G}^a\right] | \mathcal{G}^b\right].
\end{aligned}$$

We note that:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[F(X) \,|\, \mathcal{G}^a\right] | \mathcal{G}^b\right] \\
&\qquad = \int\int F(X_{A\setminus\{a,b\}}, x_a, x_b) \mathbf{P}_a((X_{A\setminus\{a,b\}}, Z), x_b, \mathrm{d}x_a) \mathbf{P}_b((X_{A\setminus\{a,b\}}, Z), \mathrm{d}x_b) \\
&\qquad = \int\int F(X_{A\setminus\{a,b\}}, x_a, x_b) \mathbb{P}^{X_b|Z}(Z, dx_b) \mathbb{P}^{X_a|Z}(Z, \mathrm{d}x_a) \\
&\qquad = \mathbb{E}\left[\mathbb{E}\left[F(X) \,|\, \mathcal{G}^b\right] | \mathcal{G}^a\right].
\end{aligned}$$

Hence, the equality follows. ☐

The key to the definition of the Malliavin framework is the so-called integration by parts.

Theorem 2.2.5 — Integration by parts I. *Let $F \in \mathcal{S}$, for every simple process U,*

$$\langle DF, U\rangle_{L^2(E_A\times A)} = \mathbb{E}\left[F\sum_{a\in A} D_a U_a\right]. \tag{2.2.3}$$

Proof of theorem 2.2.5. We get:

$$\begin{aligned}\langle DF, U\rangle_{L^2(E_A\times A)} &= \mathbb{E}\left[\sum_{a\in A} D_a F U_a\right]\\ &= \mathbb{E}\left[\sum_{a\in A}(F - \mathbb{E}\left[F\,|\,\mathcal{G}^a\right])U_a\right]\\ &= \sum_{a\in A}\mathbb{E}\left[F(U_a - \mathbb{E}\left[U_a\,|\,\mathcal{G}^a\right])\right]\\ &= \sum_{a\in A}\mathbb{E}\left[F D_a U_a\right],\end{aligned}$$

by self-adjointness of the conditional expectation. □

Corollary 2.2.6 — Closability of the discrete gradient. *The operator D is closable from $L^2(E_A)$ into $L^2(A\times E_A)$.*

Proof. The proof is analogous to the proof of closability of the gradient in (Decreusefond and Halconruy, 2019, corollary 2.5) □

The domain of D in $L^2(E_A)$ is the closure of cylindrical functionals with respect to the norm:

$$\|F\|_{1,2} := \sqrt{\|F\|^2_{L^2(E_A)} + \|DF\|^2_{A\times L^2(E_A)}}.$$

The following lemma gives a way to define square-integrable functionals in $\operatorname{Dom} D$ that are not in $\mathcal{S}$.

Lemma 2.2.7 *If there exists a sequence $(F_n)_{n\in\mathbb{N}}$ of elements of $\operatorname{Dom} D$ such that*

1. *the sequence converges to F in $L^2(E_A)$,*
2. $\sup_{n\in\mathbb{N}} \|DF_n\|_{L^2(E_A\times A)} < +\infty$,

then F belongs to $\operatorname{Dom} D$ and $DF = \lim_{n\to+\infty} DF_n$.

Proof. Let $(F_n)_{n\in\mathbb{N}}$ a sequence in $L^2(E_A)$ with $\mathbf{P}$-a.s. limit F, then for $a \in A$,

$$\begin{aligned}\mathbb{E}[|D_aF - D_aF_n|^2] &\le \mathbb{E}[|F - F_n|^2] + \mathbb{E}\left[|\mathbb{E}\left[F_n\,|\,\mathcal{G}^a\right] - \mathbb{E}\left[F\,|\,\mathcal{G}^a\right]|^2\right]\\ &\le \mathbb{E}[|F - F_n|^2] + \mathbb{E}\left[\mathbb{E}\left[|F - F_n|^2\,|\,\mathcal{G}^a\right]\right] \text{ by Jensen's inequality}\\ &= 2\mathbb{E}[|F - F_n|^2] \xrightarrow{n\to+\infty} 0.\end{aligned}$$

Let $(A_m)_{m\in\mathbb{N}}$ a family of subsets of A such that $\bigcup_{m\ge 0} A_m = A$ and $|A_m| = m$, then for all $m \in \mathbb{N}$, $(\sum_{a\in A_m} D_aF_n)_{n\in\mathbb{N}}$ converges in $L^2(E_A)$ to $\sum_{a\in A_m} D_aF$. We denote by D^m the operator on $L^2(E_A \times A)$ such that for $a \in A_m$, $D_a^m = D_a$ and otherwise D_a^m is the null operator. For $m \in \mathbb{N}$, $(D^m F_n)_{n\in\mathbb{N}}$ converges to $D^m F$ in $L^2(E_A \times A)$. Because of (2), by the uniform boundedness principle, DF is in $L^2(E_A \times A)$, and the result follows. □

Definition 2.2.8 (Divergence operator)**.** The domain of the divergence operator $\operatorname{Dom}\delta$ in $L^2(E_A)$ is the set of processes U in $L^2(E_A \times A)$ such that there exists δU satisfying the duality relation

$$\langle DF, U\rangle_{L^2(E_A\times A)} = \mathbb{E}[F\delta U], \text{ for all } F \in \operatorname{Dom} D. \tag{2.2.4}$$

Moreover, for any process U belonging to $\operatorname{Dom}\delta$, δU is the unique element of $L^2(E_A)$ characterized by that identity. The integration by parts formula entails that for every process $U \in \operatorname{Dom}\delta$,

$$\delta = \sum_{a\in A} D_a U_a. \tag{2.2.5}$$

Definition 2.2.9 (Ornstein-Uhlenbeck operator)**.** The Ornstein-Uhlenbeck operator, denoted by L is defined on its domain

$$\operatorname{Dom}\mathsf{L} = \left\{ F \in L^2(E_A) : \mathbb{E}\left[\left|\sum_{a\in A} D_a F\right|^2\right] < +\infty \right\} \supseteq \mathcal{S}$$

by

$$\mathsf{L}F := -\delta DF = -\sum_{a\in A} D_a F. \tag{2.2.6}$$

2.3 Chaos decomposition

The lemma 2.2.4 entails a chaos decomposition of $L^2(E_A)$ similar to the one in Duerinckx (2021).

Theorem 2.3.1 — Chaos decomposition. *For any $F \in L^2(E_A)$,*

$$F = \mathbb{E}\left[F\,|\,Z\right] + \sum_{n=1}^{+\infty} \pi_n(F), \tag{2.3.1}$$

where $(\pi_n)_{n\in\mathbb{N}}$ is a sequence of orthogonal projectors on $L^2(E_A)$.

Proof. One can notice that:

$$\mathbb{E}[D_a F(X)|\mathcal{G}^a] = D_a(\mathbb{E}[F|\mathcal{G}^a])F(X) = 0, \text{ for all } a \in A. \tag{2.3.2}$$

Let $(A_m)_{m\in\mathbb{N}}$ a family of finite subsets of A such that $|A_m| = m$ and $\bigcup_{m\in\mathbb{N}} A_m = A$. Let $m \in \mathbb{N}$, $\operatorname{Id}_{L^2(E_{A_m})} = \prod_{a\in A_m}(D_a + \mathbb{E}[\cdot|\mathcal{G}^a])$. Indeed, for all $a \in A_m$, $\operatorname{Id}_{\operatorname{Dom} D} = D_a + \mathbb{E}[\cdot|\mathcal{G}^a]$. Hence, by distributivity and by using lemma 2.2.4, the identity also reads off: $\operatorname{Id}_{L^2(E_{A_m})} = \sum_{n=0}^{m} \pi_n^m$, where

$$\pi_n^m := \sum_{J\subset A_m,\,|J|=n} \left(\prod_{b\in J} D_b\right)\left(\prod_{c\in A_m\setminus J} \mathbb{E}[\cdot|\mathcal{G}^c]\right) \quad \forall n \le m. \tag{2.3.3}$$

Let $n \le m$,

$$\begin{aligned}
\pi_n^m\pi_n^m &= \sum_{\substack{I\subset A_m\\|I|=n}}\sum_{\substack{J\subset A_m\\|J|=n}} \left(\prod_{b\in I} D_b\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\left(\prod_{d\in J} D_d\right)\left(\prod_{e\in A_m\setminus J} \mathbb{E}[\cdot|\mathcal{G}^e]\right)\\
&= \sum_{I\subset A_m,\,|I|=n}\ \sum_{J\subset A_m,\,|J|=n} \left(\prod_{b\in I} D_b \prod_{e\in A_m\setminus J} \mathbb{E}[\cdot|\mathcal{G}^e]\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\prod_{d\in J} D_d\right)\\
&= \sum_{\substack{I\subset A_m\\|I|=n}} \left(\prod_{b\in I} D_b \prod_{e\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\prod_{d\in I} D_d\right) \quad \text{by lemma 2.2.4}\\
&= \sum_{I\subset A_m,\,|I|=n} \left(\prod_{b\in I}\prod_{b\in I} D_b D_b\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\mathbb{E}[\cdot|\mathcal{G}^c]\right) = \pi_n^m.
\end{aligned} \tag{2.3.4}$$

By convention $\pi_n^m(F) = 0$ for $n > m$. Analogously, for $n' \neq n$, $\pi_n^m\pi_{n'}^m = 0$ The operator π_n^m is continuous on $L^2(E_A)$. Hence, $(\pi_n^m)_{m\in\mathbb{N}}$ is a well-defined family of projectors on $L^2(E_A)$. Moreover, for all $n \in \mathbb{N}$ and $F \in L^2(E_A)$, we have $\sup_{m\in\mathbb{N}} \|\pi_n^m(F)\|_{L^2(E_A)} \le \|F\|_{L^2(E_A)}$. Then, by the uniform boundedness principle,

$$\sup_{\substack{m\in\mathbb{N}\\ \|F\|_{L^2(E_A)}}} \|\pi_n^m(F)\|_{L^2(E_A)} < +\infty.$$

The pointwise limits of $(\pi_n^m(F))_{m\in\mathbb{N}}$ for $F \in L^2(E_A)$ define a bounded linear operator π_n on $L^2(E_A)$ for $n \in \mathbb{N}$. Thus:

$$L^2(E_A) = \bigoplus_{n=0}^{+\infty} \operatorname{Im} \pi_n. \tag{2.3.5}$$

Given (2.3.3), for a functional $F \in \operatorname{Dom} \mathsf{L}$, we have $\pi_0(F) = \mathbb{E}[F|Z]$. □

Lemma 2.3.2 — Spectral decomposition. *Let $F \in L^2(E_A)$ of chaos decomposition*

$$F = \mathbb{E}[F|Z] + \sum_{n=1}^{+\infty} \pi_n(F).$$

1. *We say that F belongs to* $\operatorname{Dom} \mathsf{L}$ *whenever*

$$\sum_{n=1}^{+\infty} n^2 \|\pi_n(F)\|_{L^2(E_A)} < +\infty.$$

2. The operator has a unit spectral gap, i.e. the spectrum of L *coincides with* $\mathbb{N}_0$.

$$L^2(E_A) = \bigoplus_{k=0}^{+\infty} \ker(\mathsf{L} + k\mathrm{Id}). \tag{2.3.6}$$

3. It is invertible from $L_0^2(E_A) = \{F \in L^2(E_A),\, \mathbb{E}[F|Z] = 0\}$ *into itself.*

Proof of lemma 2.3.2. Let us show that π_n is in the domain of L for all $n \in \mathbb{N}$. By summability,

$$\begin{aligned} |\sum_{a\in A} D_a\pi_n|^2 &= \left|\sum_{a\in A} D_a \sum_{I\subset A,\,|I|=n} \left(\prod_{b\in I} D_b\right)\left(\prod_{c\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\right|^2 \\ &= \left|\sum_{a\in A} \mathbb{1}_I(a) \sum_{I\subset A,\,|I|=n} \left(\prod_{b\in I} D_b\right)\left(\prod_{c\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\right|^2 \\ &= n^2 \left|\sum_{I\subset A,\,|I|=n} \left(\prod_{b\in I} D_b\right)\left(\prod_{c\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\right|^2 \quad \text{since } |I| = n \\ &= n^2|\pi_n|^2, \end{aligned} \tag{2.3.7}$$

so for $F \in L^2(E_A)$, $\pi_n(F) \in \mathrm{Dom}\,\mathsf{L}$. Hence, because of the orthogonality of $(\mathrm{Im}\ \pi_n)_{n\in\mathbb{N}}$, $F \in \mathrm{Dom}\, L \iff \sum_{n=1}^{+\infty} n^2\|\pi_n(F)\|_{L^2(E_A)} < +\infty$. With the same calculations, we get $\mathsf{L}\pi_n = -n\pi_n$. The spectrum of $-\mathsf{L}$ coincides with $\mathbb{N}$. Then, we deduce that:

$$\mathsf{L} = \sum_{n=0}^{+\infty} -n\pi_n, \tag{2.3.8}$$

and $\mathrm{Im}\,\pi_n \subset \ker(\mathsf{L} + \mathsf{n}\mathrm{Id})$. Because of the orthogonality of the kernels, we get $\mathrm{Im}\,\pi_n = \ker(\mathsf{L} + \mathsf{n}\mathrm{Id})$. Now let us prove the third item. The pseudoinverse L^{-1} is defined on its domain $\{F \in L^2(E_A) : \mathbb{E}[F|Z] = 0\}$ and reads $\sum_{n=1}^{+\infty} -\frac{\pi_n}{n}$. Then for $F \in \{G \in \mathrm{Dom}\,\mathsf{L} : \mathbb{E}[G|Z] = 0\}, \mathsf{L}^{-1}(\mathsf{L}F) = F$. □

Corollary 2.3.3 *For* $k > 0$ *and* J *a subset of* A *of cardinal* k, *let us denote by* $\mathfrak{C}_k$ *the space of functionals* $\phi = \sum_{J\subset A, |J|=k} \psi_J$ *such that:*

- *for every* $J \subset A$, ψ_J *is* $\mathcal{F}_J$*-measurable;*
- *for every* $K \subset A$, $\mathbb{E}[\psi_J|\mathcal{G}_K] = 0$ *unless* $K \subset J$;

then $\mathfrak{C}_k = \ker(\mathsf{L} + k\mathrm{Id})$.

Proof of corollary 2.3.3. From (2.3.3), for $J = (a_1, \ldots, a_n) \subset A$, the component ψ_J is $\mathcal{F}_J$-measurable. Let us compute the expression of the iterated gradient for F a $\mathcal{F}_J$-measurable function:

$$\prod_{a\in J} D_a F = \sum_{k=0}^{|J|}(-1)^k \sum_{\substack{K\subseteq J\\|K|=k}} \mathbb{E}\left[F\,|\,\mathcal{G}^K\right]$$
$$= \sum_{L\subseteq J}(-1)^{|J|-|L|}\mathbb{E}\left[F\,|\,\mathcal{G}_L\right],$$

where $\mathcal{G}^K = \sigma(X^K)\vee\sigma(Z)$ and $\mathcal{G}_L = \sigma(X_L)\vee\sigma(Z)$. In this view, we have the inclusion $\ker(\mathsf{L}+n\mathrm{Id}) = \mathrm{Im}\,\pi_n \subset \mathfrak{C}_n$ for $n\in\mathbb{N}$.
Conversely, let ϕ for which the properties above hold.

$$\begin{aligned}\mathsf{L}\phi &= -\sum_{a\in A} D_a \sum_{J\subset A,|J|=n}\psi_J\\ &= -\sum_{a\in A}\sum_{J\subset A,|J|=n}\left(\psi_J - \mathbb{E}\left[\psi_J\,|\,\mathcal{G}^a\right]\right)\\ &= -\sum_{k\in A}\sum_{\substack{J\subset A,|J|=n\\ a\in J}}\psi_J \text{ because } \mathbb{E}\left[\psi_J\,|\,\mathcal{F}_{A\setminus\{a\}}\right]=0 \text{ for } J\not\subset A\setminus\{a\}\\ &= -n\sum_{J\in A,|J|=n}\psi_J = -n\phi.\end{aligned}$$

Therefore, $\mathfrak{C}_n = \ker(\mathsf{L}+n\mathrm{Id})$ for $n\geq 1$. □

Remark 2.3.4. It is known that there is no Hoeffding decomposition for functionals of exchangeable random variables unless we assume *weak independence* (Peccati, 2004). The decomposition at hand is not a *Hoeffding decomposition* since the kernels of the U-statistics also depend on Z. The first term in the decomposition is not $\mathbb{E}[F]$ as expected.

2.4 Dirichlet structure

The map L can be viewed as the generator of a Glauber dynamics where the index set is a finite set of random variables indexed by A_m for $m>1$. For practical term, we introduce a new index ∂ and $X_\partial = Z$ $\mathbb{P}$-a.s.

Definition 2.4.1 (Modified Glauber process)**.** Consider $(N(t))_{t\geq 0}$ a Poisson process on the half-line $[0,+\infty)$ of rate $|A_m|+1$. Let $(X^{\circ A_m}(t))_{t\geq 0} = (X_a^{\circ A_m}(t), t\geq 0, a\in A)$ the process valued in E_A starting with $X^{\circ A_m}(0) = X$ which evolves according to the following rule. At jump time τ of the process,

- Choose randomly an index a in $A_m \sqcup \{\partial\}$ with equal probability.
- If $a\neq\partial$, replace $X_a^{\circ A_m}(\tau)$ with a conditionally independent random variable $X_a^{`}$ distributed according to $\mathbf{P}_a((X_{A\setminus\{a\}}^{\circ A_m}(\tau), Z),\cdot)$, otherwise do nothing.

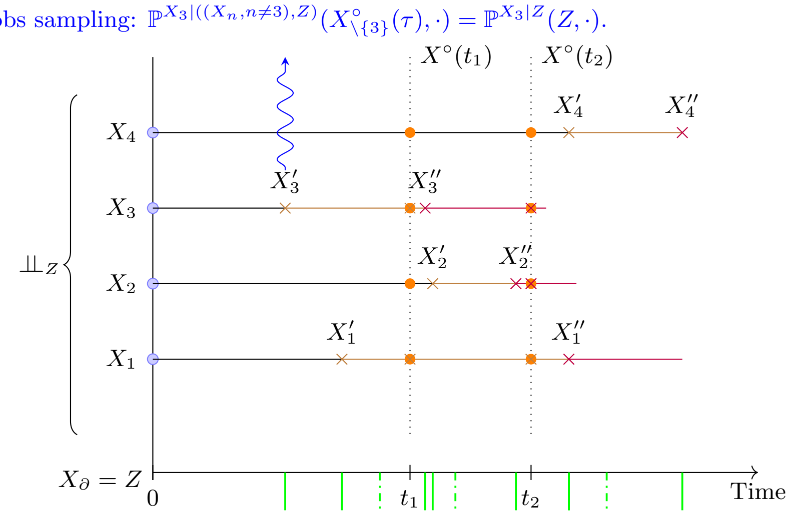


Figure 2.1: Modified Glauber dynamics

That Markov process has for infinitesimal generator L^{A_m}:

$$\mathsf{L}^{A_m} F = - \sum_{a \in A_m} D_a F.$$

That is referred as Glauber dynamics because one can identify a coordinate, and change the random variable at this coordinate in function of the other ones. Here it does not simulate the Ising model, but a more general model. It is mostly known for its application to spin system as a particular case that involves a probability measure on a product space (Sambale and Sinulis, 2020). These finite spin systems can have a rich dependence structure among the sites. The algorithm gives us the construction of the Markov process. Our description is similar to the Glauber dynamics described in Adamczak et al. (2022, section 4.3.). Our aim is to show that the operator L is an infinitesimal generator, letting $m \to +\infty$. We recall the Hille-Yosida theorem (Yosida, 1995).

Proposition 2.4.2 — Hille-Yosida. *A linear operator L on $L^2(E_A)$ is the generator of a strongly continuous contraction semigroup on $L^2(E_A)$ if and only if*

1. $\operatorname{Dom} L$ *is dense in* $L^2(E_A)$*;*

2. L *is dissipative, i.e. for any* $\lambda > 0$, $F \in \operatorname{Dom} L$,

$$\|\lambda F - LF\|_{L^2(E_A)} \geq \lambda \|F\|_{L^2(E_A)}.$$

3. $\operatorname{Im}(\lambda Id - L)$ *is dense in* $L^2(E_A)$.

Theorem 2.4.3 L *is an infinitesimal generator on E_A of a strongly continuous contraction semigroup on $L^2(E_A)$.*

Proof of theorem 2.4.3. We know that $\mathcal{S}$ is dense in $L^2(E_A)$. As $\operatorname{Dom} \mathsf{L} \supset \mathcal{S}$, it is also dense in $L^2(E_A)$. Let A_m an increasing sequence (with respect to $\subset$) of subsets of A such that

$\bigcup_{n\geq 1} A_m = A \cup \partial$ and $|A_m| = m$. Then $(\mathcal{F}_{A_m})_{m\in\mathbb{N}}$ is a filtration. For $F \in L^2(E_A)$, let $F_m = \mathbb{E}\left[F | \mathcal{F}_{A_m}\right]$. Since $(F_m)_{m\in\mathbb{N}}$ is a square-integrable $\mathcal{F}_A$-martingale, $(F_m)_{m\in\mathbb{N}}$ converges both almost surely and in $L^2(E_A)$ to F. For any $m \in \mathbb{N}$, F_m depends only on X_{A_m}. Because of the conditional independence of the random variables X_a given X_∂, for all $a \in A$, we get that $D_a F_m = \mathbb{E}\left[D_a F | \mathcal{F}_{A_m}\right]$. Using that L_{A_m} is dissipative for all $m \in \mathbb{N}$, we have:

$$\lambda^2 \|F_m\|^2_{L^2(E_A)} \leq \|\lambda F_m - \mathsf{L}^{A_m} F_m\|^2_{L^2(E_A)} = \mathbb{E}\left[\left(\lambda F_m + \sum_{a\in A_n} D_a F_m\right)^2\right]$$

$$= \mathbb{E}\left[\left(\lambda F_m + \sum_{a\in A} D_a F_m\right)^2\right] \text{ because } D_a F_m = 0, \text{ if } a \notin A_m.$$

$$= \mathbb{E}\left[\mathbb{E}\left[\lambda F + \sum_{a\in A} D_a F \,\middle|\, \mathcal{F}_{A_m}\right]^2\right].$$

It means that the operator L is dissipative. Thus, by the Hille-Yosida theorem, L is the infinitesimal generator of a strongly continuous contraction semigroup on $L^2(E_A)$ denoted P. □

Lemma 2.4.4 *Let $F \in L^2(E_A)$, then*

$$\mathbb{E}\left[F(X^{\circ A_m}) | X, Z\right] = P_t^{A_m} F \xrightarrow{\mathbf{P}-a.s.} P_t F$$

and

$$X^{\circ A_m} \xrightarrow{d} X^\circ.$$

Proof of lemma 2.4.4. The theorem 17.25 of (Kallenberg, 1997, Trotter, Sova, Kurtz, Mackevičius) gives the convergence in distribution of $X^{\circ A_m}$ towards X° the Markov process associated to L, and the almost sure convergence of the semigroup. □

These are pieces of the Dirichlet structure with invariant measure $\mathbf{P}$ that we complete with the carré du champ operator. Here, we note that $\mathcal{S}$ is an algebra which is a core of $\operatorname{Dom} \mathsf{L}$.
By an argument of density, there exists an algebra $\mathcal{A} \supset \mathcal{S}$ maximal in the sense of inclusion such that the carré du champ operator acts on it.

Definition 2.4.5 (Dirichlet structure)**.** The associated Dirichlet structure defined on $(E_A, \Upsilon, \mathbf{P})$ is given by the quadruple $(X^\circ, \mathsf{L}, (P_t)_{t\geq 0}, \mathcal{E})$ where X° is a Markov process with values in E_A whose infinitesimal generator is L and its semigroup is P, i.e. for any $F \in L^\infty(E_A)$:

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t F = (\mathsf{L} P_t) F.$$

Furthermore, $\mathbf{P}^Z$ is the invariant (or stationary) distribution of X° given Z and the Dirichlet form is defined by

$$\mathcal{E}(F, G) = \mathbb{E}[\Gamma(F, G)].$$

It comes with the classical properties entailed by the spectral decomposition of L, including the Mehler's formula.

Lemma 2.4.6 — Mehler's formula. *For any* $F \in L^2(E_A)$*,*

1.
$$\begin{aligned} P_t F &= \mathbb{E}[F \mid Z] + \sum_{n=1}^{\infty} e^{-nt} \pi_n(F) \\ &= \mathbb{E}[F(X^\circ(t)) \mid X], \end{aligned} \tag{2.4.1}$$

In particular $P_t F \in \operatorname{Dom} \mathsf{L} \cap \operatorname{Dom} \mathsf{L}^{-1}$.

2.
$$\lim_{t \to \infty} P_t F(X) = \mathbb{E}[F(X) \mid Z]$$

3. *The pseudoinverse of* L *can be written:*
$$\mathsf{L}^{-1} F := -\int_0^{+\infty} P_t F \, \mathrm{d}t.$$

Proof of lemma 2.4.6. Since formally $P_t = e^{-t\mathsf{L}}$, we get the first line of (2.4.1) from the spectral decomposition of L. The second line is deduced from the definition of the Glauber dynamics and by passing to the limit. Then,

$$\begin{aligned} \mathbb{E}[F \mid Z] - F &= \lim_{t \to +\infty} P_t F - P_0 F \\ &= \int_0^{+\infty} \frac{\mathrm{d}}{\mathrm{d}t} P_t F \, \mathrm{d}t \\ &= \mathsf{L}\left(\int_0^{+\infty} P_t F \, \mathrm{d}t \right). \end{aligned}$$

Taking $\mathbb{E}[F \mid Z] = 0$, we get the expression of the pseudoinverse. □

Remark 2.4.7. By the chaos expansion, $P_t F$ can be defined as the limit in $L^2(E_A)$ of elements $(P_t F_n)_{n \in \mathbb{N}}$ for F_n in $\mathcal{S}$. Hence, it is sufficient to define the semigroup acting on a functional of some finite vector of random variables X_B, using the definition of the Glauber dynamics entailed by it.

The infinitesimal generator satisfies another integration by parts formula due to the Dirichlet structure which is the key to investigating the so-called fourth moment phenomenon.

Lemma 2.4.8 — Integration by parts II. *For* $(F, G) \in \mathcal{A}^2$*,*
$$\mathcal{E}(F, G) = -\mathbb{E}[F \mathsf{L} G]. \tag{2.4.2}$$

We introduce to the difference operator which is associated to the Malliavin-Dirichlet structure at hand. That difference operator serves the same purpose as in Lachièze-Rey et al. (2017) and Dung (2021) for computations in the proofs of the limit theorems.

Definition 2.4.9 (Difference operator)**.** Let $F : E_A \to \mathbb{R}$, for $a \in A$, we introduce the operator

$$\Delta^{\{a\}} F : E_A \times E_a \longrightarrow \mathbb{R}$$
$$(x, x'_a) \longmapsto f(x) - f(x^{\{a\}}, x'_a).$$

For the sake of conciseness, we shall write $F^{\{a\}'} = F(X^{\{a\}}, X'_a)$.

Lemma 2.4.10 *For F a functional in* Dom D, *the gradient also reads as:*

$$D_a F = \mathbb{E}\left[\Delta^{\{a\}} F(X, X'_a) | X, Z\right], \tag{2.4.3}$$

where X'_a has the law of X_a given Z and is conditionally independent of $X^{\{a\}}$ given Z. Similarly,

$$\Gamma(F, G) = \frac{1}{2} \sum_{a \in A} \mathbb{E}\left[\left(\Delta^{\{a\}} F(X, X'_a)\right)\left(\Delta^{\{a\}} G(X, X'_a)\right) | X, Z\right]. \tag{2.4.4}$$

Proof. We have

$$\mathbb{E}\left[F \,|\, \mathcal{G}^a\right] = \int F(X^{\{a\}}, x_a) \mathbf{P}_a(\,\mathrm{d}x_a).$$

Since $\sigma(X_a)$ is independent of $\sigma(X^{\{a\}})$ given $\sigma(Z)$, we obtain

$$\mathbb{E}\left[F \,|\, \mathcal{G}^a\right] = \int F(X^{\{a\}}, x_a) \mathbb{P}^{X_a|Z}(\,\mathrm{d}x_a).$$

Eqn.(2.4.4) is proved similarly. □

2.5 Functionals identities

This section is devoted to classical functional identities obtained in the Malliavin framework. We follow the approach of Houdré et al. (2002) using a covariance identity based on difference operators to deduce concentration inequalities.

Proposition 2.5.1 *For $F \in L^2(E_A)$ and $a \in A$, then:*

$$D_a(P_t F) = e^{-t} \mathbb{E}\left[\Delta^{\{a\}} F(X^{\circ}(t), X'_a) \,|\, X, Z\right] \tag{2.5.1}$$

where X' has the law of X given Z.

Proof of proposition 2.5.1. We consider the Glauber dynamics with index set a finite subset A_m of A, as the construction of process $(X^{\circ A_m}(t))_{t \in \mathbb{R}^+}$ is explicit in that case. Let $a \in A_m$, we denote by N_a the Poisson process of intensity 1 which represents the life duration of the a-th component in the dynamics of $X^{\circ A_m}(t)$, so:

$$X_a^{\circ A_m}(t) = \mathbb{1}_{\{\tau_a \geq t\}} X_a + \mathbb{1}_{\{\tau_a < t\}} X'_a,$$

where $\tau_a = \inf\{t \geq 0,\, N_a(t) \neq N_a(0)\}$ is the life duration of the a-the component of the original sequence, exponentially distributed with parameter 1 (independent of everything else) and X'_a

is conditionally independent of X given Z. Then:

$$\begin{aligned} D_a P_t^{A_m} F &= P_t^{A_m} F - \mathbb{E}\left[P_t^{A_m} F | \mathcal{G}_a\right] \\ &= P_t^{A_m} F - \mathbb{E}\left[\mathbb{E}\left[F(X^{\circ A_m}(t))|X, Z\right] \mathbb{1}_{\{t \le \tau_a\}} | \mathcal{G}_a\right] - \mathbb{E}\left[F(X^{\circ A_m}(t))\mathbb{1}_{\{t > \tau_a\}}|X, Z\right] \\ &= \mathbb{E}\left[F(X^{\circ A_m}(t))\mathbb{1}_{\{t \le \tau_a\}}|X, Z\right] - \mathbb{E}\left[\mathbb{E}\left[F(X^{\circ A_m}(t))|X, Z\right] \mathbb{1}_{\{t \le \tau_a\}} | \mathcal{G}_a\right] \\ &= e^{-t}\mathbb{E}\left[\Delta^{\{a\}} F(X^{\circ A_m}(t), X'_a) | X, Z\right] \end{aligned}$$

because the law of $X^{\backprime}_a$ given X is the same as the one of X'_a given X.
On one hand,

$$D_a P_t^{A_m} F \xrightarrow{\mathbb{P}-a.s.} D_a P_t F.$$

On the other hand, by the Skorohod's representation theorem, there exist copies of $X^{\circ A_m}$ and X° on a common probability space $(\tilde{\Omega}, \tilde{\mathcal{T}}, \tilde{\mathbb{P}})$ such that the sequence $(X^{\circ A_m})_{m \in \mathbb{N}}$ converges to X° $\tilde{\mathbb{P}}$-a.s. As the whole structure is invariant by copy, we can suppose the almost sure convergence on $(\Omega, \mathcal{T}, \mathbb{P})$, and the relation passes to the limit. □

Remark 2.5.2. In the case, we have only one random variable (or one particle), then the commutation relation simplifies to $D_a(P_t F) = D_a$.

Corollary 2.5.3 — Conditional covariance identity. *For any $F, G \in L^2(E_A)$, then:*

$$\operatorname{Cov}(F, G|Z) = \int_0^\infty e^{-t} \sum_{a \in A} \mathbb{E}\left[(D_a F)(\Delta^{\{a\}} G(X^{\circ}(t), X'_a)) | Z\right] \mathrm{d}t. \tag{2.5.2}$$

Proof of corollary 2.5.3. We use the following conditional covariance formula analogous to the covariance formula:

$$\operatorname{Cov}(F, G|Z) = \mathbb{E}\left[FG|Z\right] = \mathbb{E}\left[F \,\mathrm{L}\,\mathrm{L}^{-1} G|Z\right] \tag{2.5.3}$$

By the integration by parts I (2.2.3) which also holds with conditional expectation given Z, we get:

$$\begin{aligned} \mathbb{E}\left[F \,\mathrm{L}\,\mathrm{L}^{-1} G|Z\right] &= -\sum_{a \in A} \mathbb{E}\left[(D_a F)(D_a \mathrm{L}^{-1} G)|Z\right] \\ &= -\sum_{a \in A} \mathbb{E}\left[(D_a F)(D_a \int_0^\infty P_t G \,\mathrm{d}t)|Z\right] \\ &= -\sum_{a \in A} \mathbb{E}\left[(D_a F)\left(\int_0^\infty D_a P_t G \,\mathrm{d}t\right)|Z\right] \\ &= -\int_0^\infty e^{-t} \sum_{a \in A} \mathbb{E}\left[(D_a F)\mathbb{E}\left[\Delta^{\{a\}} G(X^{\circ}(t), X'_a)|X, Z\right]|Z\right] \mathrm{d}t, \end{aligned}$$

using (2.5.1). □

As an immediate consequence of the spectral gap, we find another proof of the Efron-Stein inequality which is of independent interest.

Proposition 2.5.4 *If $F \in \mathfrak{C}_p$ then*

$$\operatorname{Var}[F] = \frac{1}{p}\mathcal{E}(F) = \frac{1}{p}\|DF\|_{L^2(E_A)}.$$

Moreover, if there exist $F_1, \ldots, F_m \in L^2(E_A)$ such that $F = \sum_{p=1}^m F_p$ with $F_p \in \mathfrak{C}_p$ for $p \in [\![1, m]\!]$, then:

$$\operatorname{Var}[F] \leq \|DF\|_{L^2(E_A)}. \tag{2.5.4}$$

Proof of proposition 2.5.4. Let us use the previous covariance identity, we have:

$$\begin{aligned}
\operatorname{Var}[F] = \operatorname{Cov}(F,F) &= \mathbb{E}[\Gamma(F, -\mathsf{L}^{-1}F)] \\
&= \mathbb{E}\left[\Gamma\left(\sum_{p=1}^m F_p, \sum_{q=1}^m \frac{1}{q}F_q\right)\right] \\
&= \sum_{p=1}^m \sum_{q=1}^m \frac{1}{q}\mathbb{E}\left[\Gamma\left(F_p, F_q\right)\right] \\
&= \sum_{p=1}^m \frac{1}{p}\mathbb{E}\left[\Gamma\left(F_p, F_p\right)\right] \text{ because } \mathbb{E}[\Gamma(F_p, F_q)] = 0 \text{ for } q \neq p
\end{aligned}$$

It yields the inequality (2.5.4) noting that $\Gamma(F_p, F_p) \geq 0$ for all $p > 0$. □

We now deduce the conditional first-order Poincaré inequality for functionals of conditionally independent random variables. The equivalent for functionals of independent random variables is rather known as the Efron-Stein inequality in the literature (Efron and Stein, 1981).

Theorem 2.5.5 — Conditional Efron-Stein inequality. *For $F \in L^2(E_A)$ such that $\mathbb{E}[F\,|\,Z] = 0$,*

$$\operatorname{Var}[F|Z] \leq \mathbb{E}\left[\Gamma(F,F)\,|\,Z\right]. \tag{2.5.5}$$

Proof of theorem 2.5.5. The conditional covariance formula yields

$$\begin{aligned}
\operatorname{Var}[F|Z] &= \int_0^\infty e^{-u} \sum_{a \in A} \mathbb{E}\left[(D_aF)(\Delta^{\{a\}}F)(X_u^\circ, X_a')\,|\,Z\right] \mathrm{d}u \\
&\leq \int_0^\infty e^{-u} \sqrt{\sum_{a \in A} \mathbb{E}[(D_aF)^2|Z]}\sqrt{\sum_{a\in A}\mathbb{E}[\mathbb{E}\left[(\Delta^{\{a\}}F)(X_u^\circ, X_a')|X,Z\right]^2|Z]}\, \mathrm{d}u.
\end{aligned}$$

The invariance of $\mathbf{P}^Z$ under the Glauber dynamics entails that

$$\sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[(\Delta^{\{a\}}F)(X_u^\circ, X_a')|X,Z\right]^2|Z\right] = \sum_{a\in A}\mathbb{E}[(D_aF)^2|Z].$$

Hence,

$$\operatorname{Var}[F|Z] \leq \mathbb{E}[\Gamma(F,F)|Z],$$

proving the theorem. □

Remark 2.5.6 (The optimal constant in the Poincaré inequality)**.** As mentioned in Bakry, Gentil, and Ledoux (2013), if a Poincaré inequality $\mathcal{P}(C)$ with constant C holds for the Dirichlet structure at hand, the spectrum of the symmetric positive operator $-\mathsf{L}$ is included in $\{0\} \cup [\frac{1}{C}, \infty]$. The existence of $\mathcal{P}(C)$ does not tell whether the spectrum of L is discrete or not. The consequence of $\mathcal{P}(C)$ is the fact that $\mathcal{E}(F) = 0 \implies F$ is constant.

We find a version of the McDiarmid's inequality for conditionally independent random variables.

Theorem 2.5.7 — Conditional McDiarmid's inequality. *Let F be a square-integrable functional such that for all $a \in A$:*

$$\sup_{\substack{x^{\{a\}} \in E_{A\setminus\{a\}} \\ x'_a \in E_a}} |F(x^{\{a\}}, x'_a) - F(x)| \leq d_a.$$

For any $x > 0$, we have the inequality:

$$\mathbb{P}(F(X) - \mathbb{E}\left[F(X)|Z\right] \geq x|Z) \leq \exp\left(-\frac{x^2}{2\sum_{a\in A} d_a^2}\right). \tag{2.5.6}$$

Our strategy of proof is different from the original McDiarmid's original proof in McDiarmid (1989).

Proof of theorem 2.5.7. We assume that $F = F(X)$ is a bounded random variable verifying $\mathbb{E}\left[F|Z\right] = 0$. Using the inequality:

$$|e^{tx} - e^{ty}| \leq \frac{t}{2}|x-y|(e^{tx} + e^{ty}) \quad \forall\, x, y \in \mathbb{R}. \tag{2.5.7}$$

We have:

$$\begin{aligned} |\Delta^{\{a\}} e^{tF}(X, X'_a)| &= |e^{tF} - e^{tF^{\{a\}'}}| \\ &\leq \frac{t}{2}|\Delta^{\{a\}} F(X, X'_a)| \left(e^{tF} + e^{tF^{\{a\}'}}\right). \end{aligned}$$

Applying the covariance identity, it yields:

$$\begin{aligned} \mathbb{E}[Fe^{tF}|Z] &= \int_0^\infty e^{-u} \sum_{a\in A} \mathbb{E}[D_a e^{tF} \Delta^{\{a\}} F(X_u^\circ, X'_a)|Z] \,\mathrm{d}u \\ &\leq \int_0^\infty e^{-u} \sum_{a\in A} \mathbb{E}\left[\mathbb{E}\left[|\Delta^{\{a\}} e^{tF}(X, X'_a)||X, Z\right] \Delta^{\{a\}} F(X_u^\circ, X'_a)|Z\right] \mathrm{d}u \\ &\leq \frac{t}{2}\int_0^\infty e^{-u} \sum_{a\in A} \mathbb{E}\left[|\Delta^{\{a\}} F(X, X'_a)|e^{tF}|\Delta^{\{a\}} F(X_u^\circ, X'_a)|Z\right] \mathrm{d}u \\ &+ \frac{t}{2}\int_0^\infty e^{-u} \sum_{a\in A} \mathbb{E}\left[|\Delta^{\{a\}} F(X, X'_a)|e^{tF^{\{a\}'}}|\Delta^{\{a\}}(X_u^\circ, X'_a)||Z\right] \mathrm{d}u \end{aligned}$$

by using the Jensen's inequality for conditional expectation in the second inequality. Since $|\Delta^{\{a\}}F(X, X'_a)|^2 \leq d_a$, $|\Delta^{\{a\}}F(X^\circ_u, X'_a)| \leq d_a$ for all $u \in \mathbb{R}^+$ and $\mathbb{E}[e^{tF^{\{a\}'}}|Z] = \mathbb{E}[e^{tF}|Z]$, this shows that:

$$\mathbb{E}[Fe^{tF}|Z] \leq \left(\sum_{a\in A} d_a^2\right) t\mathbb{E}[e^{tF}|Z] = tK^2\mathbb{E}[e^{tF}|Z],$$

where $K^2 := \sum_{a\in A} d_a^2$. Thus, in all generality for F bounded:

$$\begin{aligned}\log \mathbb{E}[e^{t(F-\mathbb{E}[F])}|Z] &= \int_0^t \frac{\mathbb{E}[(F-\mathbb{E}[F|Z])e^{s(F-\mathbb{E}[F|Z])}|Z]}{\mathbb{E}[e^{s(F-\mathbb{E}[F])}]}\,\mathrm{d}s \\ &\leq K^2 \int_0^t s\ \mathrm{d}s = \frac{t^2}{2}K^2,\end{aligned}$$

hence:

$$\begin{aligned}e^{tx}\mathbb{P}(F-\mathbb{E}[F|Z] > x|Z) &\leq \mathbb{E}[e^{t(F-\mathbb{E}[F|Z])}|Z] \\ &= e^{t^2K^2/2}, \quad t \geq 0,\end{aligned}$$

and:

$$\mathbb{P}(F-\mathbb{E}[F|Z] \geq x|Z) \leq e^{\frac{t^2}{2}K^2 - tx}, \quad t \geq 0.$$

The minimum of the right-hand side is obtained for $t = x/K^2$. If F is not bounded, the conclusion holds for $F_n = \max(-n, \min(F, n))$, $n \geq 0$, and $(F_n)_{n\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. to F. Hence:

$$\mathbb{P}(F-\mathbb{E}[F|Z] \geq x|Z) \leq \exp\left(-\frac{x^2}{2K^2}\right) = \exp\left(-\frac{x^2}{2\sum_{a\in A} d_a^2}\right).$$

The proof is thus complete. □

Remark 2.5.8. The McDiarmid's inequality for conditionally independent random variables can be recovered from Azuma's inequality. Malliavin calculus offers an alternative for proof of concentration results to classical inequalities that use either or the martingale.

Example 2.5.9. See Raginsky and Sason (2014) for applications of such logarithm Sobolev inequality.

We do not have a second-order Poincaré inequalities (Nourdin et al., 2009), because of the absence of chain rule. However, we can still proceed with the Stein's method in the next section.

2.6 Applications to normal approximation

We deal with the problem of approximations in law for functionals of the form:

$$F(X) := F(X_1, X_2, \ldots). \tag{2.6.1}$$

When the functional only depends on a random vector $(X_1, \dots, X_n)$ for a given $n \in \mathbb{N}$, the normal approximation of $F(X_1, X_2, ..., X_n)$ has been studied by Chatterjee et al. (2008); Lachièze-Rey et al. (2017). More recently, those methods have been adapted to the infinite case when $n \to \infty$ (Privault and Serafin, 2018; Dung, 2018, 2019; Arras et al., 2019; Decreusefond and Halconruy, 2019; Duerinckx, 2021; Privault and Serafin, 2022), explicitly stating quantitative limit theorems. They use Malliavin calculus for integration by parts in Stein's method. Earlier, Chen et al. (2007) derived Berry-Esseen bound for nonlinear statistics such as U-statistics and L-statistics with an order of magnitude for the bounds which corresponds to the normalizing rate of central limit theorems. They used a concentration inequality approach by Stein's method as to control the non-linear part of the statistics. Still, the derived bounds were non-uniform except for particular cases, and the sequence requires to be independent identically distributed. The combination of Stein's method and Malliavin calculus leads to finer-grained upper bounds in probability distance for limit theorems of non-linear U-statistics, essentially relying on a *chaos expansion* of square-integrable functionals.

2.6.1 Bounds in probability distance

Remove the definition of the distance defined in the previous chapter

The goal is to bound for instance the 1-Wasserstein distance

$$d_W(\mathcal{L}(F(X)), \mathcal{L}(Y))| := \sup_{h \in \mathcal{H}} |\mathbb{E}[h(F(X))] - \mathbb{E}[h(Y)]|$$

for $\mathcal{H}$ the set of 1-Lipschitz functions and Y the random variable following the target distribution. We recall the lemma 4.2 of Chatterjee et al. (2008) which provides with a standard implementation of the Stein's method for this probabilistic distance with respect to the normal distribution $\mathcal{N}(0,1)$.

Lemma 2.6.1 — Normal approximation. *Let $L^\dagger h(x) := h'(x) - xh(x)$. Then,*

$$d_W(\mathcal{L}(F(X)), \mathcal{N}(0,1)) \le \sup_{\varphi \in \mathcal{H}_*} \left| \mathbb{E}[L^\dagger \varphi(F(X))] \right|, \tag{2.6.2}$$

where $\mathcal{H}_* := \{h \in C^2(\mathbb{R}, \mathbb{R}) : \ \|h'\|_\infty \le \sqrt{\frac{2}{\pi}}, \|h''\|_\infty \le 2\}$

In the following, we denote $d_W(\mathcal{L}(F(X)), \mathcal{N}(0,1))$ by $d_W(F, \mathcal{N}(0,1))$. For sake of conciseness, we denote by $\Delta^{\{a\}'}F$ the quantity $\Delta^{\{a\}}F(X, X'_a)$.

2.6.2 Rates in Lyapunov's conditional central limit

Lemma 2.6.2 *For any $F \in \mathcal{S}$ such that $\mathbb{E}[F|Z] = 0$. Then,*

$$d_W(F, \mathcal{N}(0,1)) \le \sup_{\psi \in \mathcal{H}_*} \left| \mathbb{E}\left[\sum_{a \in A} \psi(F(X^{\{a\}}, X'_a)) \Delta^{\{a\}'} F D_a(-\mathsf{L}^{-1}F) - \psi(F) \right] \right| + \sum_{a \in A} \mathbb{E}[(\Delta^{\{a\}'}F)^2 |D_a \mathsf{L}^{-1} F|]. \tag{2.6.3}$$

Proof of lemma 2.6.2. We compute:

$$\sup_{f^\dagger \in \mathcal{H}_*} |\mathbb{E}[F(f^\dagger)(F) - (f^\dagger)'(F)]|.$$

Since F is centered,

$$\begin{aligned}\mathbb{E}[F(f^\dagger)(F)] &= \mathbb{E}[\mathsf{L}(\mathsf{L}^{-1}F)f^\dagger(F)] \\ &= -\sum_{a\in A}\mathbb{E}[D_a\mathsf{L}^{-1}FD_af^\dagger(F)] \text{ by integration by parts} \\ &= -\sum_{a\in A}\mathbb{E}\left[D_a\mathsf{L}^{-1}F\mathbb{E}\left[(f^\dagger)'(F)-f^\dagger(F^{\{a\}'})|X,Z\right]\right] \\ &= -\sum_{a\in A}\mathbb{E}[D_a\mathsf{L}^{-1}F\Delta^{\{a\}'}f^\dagger(F)]\end{aligned}$$

Then, we use the Taylor expansion taking the reference point to be $F^{\{a\}'}$ instead of F, for all $a \in A$ yielding:

$$\begin{aligned}\Delta^{\{a\}'}f^\dagger(F) &= f^\dagger(F) - f^\dagger(F^{\{a\}'}) \\ &= (f^\dagger)'(F^{\{a\}'})\Delta^{\{a\}'}F + R_a,\end{aligned}$$

with $|R_a| \le \frac{\|(f^\dagger)''\|_\infty}{2}(\Delta^{\{a\}'}F)^2 = (\Delta^{\{a\}'}F)^2$. Then,

$$\begin{aligned}&|\mathbb{E}[Ff^\dagger(F) - (f^\dagger)'(F)]| \\ &\qquad\le \left|\mathbb{E}\left[\sum_{a\in A}\Delta^{\{a\}'}F(D_a(-\mathsf{L}^{-1}F))\left((f^\dagger)'(F^{\{a\}'}) - (f^\dagger)'(F)\right)\right]\right| \\ &\qquad+ \sum_{a\in A}\mathbb{E}[(\Delta^{\{a\}'}F)^2|D_a\mathsf{L}^{-1}F|].\end{aligned}$$

Because $(f^\dagger)''$ has Lipschitz-constant equal to 2, we get the result. □

We prove a quantitative Lyapunov's conditional central limit theorem for random variables with moments of order 3.

Corollary 2.6.3 — Lyapunov's conditional central limit theorem. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of thrice integrable, conditionally independent random variables given a latent random variable Z. Let us observe that*

$$\sigma_{j,Z}^2 = \mathrm{Var}(X_j|Z),\ s_{n,Z}^2 = \sum_{j=1}^n \sigma_{j,Z}^2 \text{ and } \bar{X}_n = \frac{1}{s_{n,Z}}\sum_{j=1}^n (X_j - \mathbb{E}[X_j|Z]).$$

Then,

$$d_W(\bar{X}_n, \mathcal{N}(0,1)) \le 2(\sqrt{2}+1)\mathbb{E}\left[\frac{1}{s_{n,Z}^3}\sum_{i=1}^n |X_i - \mathbb{E}[X_i|Z]|^3\right]. \tag{2.6.4}$$

The proof of the corollary follows the same steps as the one of (Decreusefond and Halconruy, 2019, Corollary 5.11), using lemma 2.6.2.
A version of conditional central limit theorem was first stated in Rao (2009) without proof and then was proved in Grzenda and Zieba (2008) for sequence of conditionally independent random variables (see Yuan et al. (2014); Bulinski (2017) for more "stringent" proofs). They also give a Lindebergh central limit theorem. We rule out a Lyapunov's conditional central limit theorem, giving a quantitative limit theorem. Let us note that our approach differs from the use of Stein's method for conditional central limit in Dey and Terlov (2023).

Example 2.6.4 (Conditional Bernoulli random variables)**.** Let $(U_i)_{i\in\mathbb{N}}$ independent uniform random variables, and $X_i = \mathbb{1}_{\{U_i \le Z\}}$, with Z an arbitrary random variable lying in $[0,1]$, then $(X_i)_{i\in\mathbb{N}}$ forms a sequence of conditionally independent random variables given Z. The law of $\mathcal{L}(X_i|X^{\{i\}}, Z)$ is a Bernoulli law of parameter Z. We compute the right-hand side of the Lyapunov theorem in this case.

$$s_{n,Z}^2 = nZ(1-Z)$$
$$\mathbb{E}\left[|X_i - \mathbb{E}\left[X_i|Z\right]|^3|Z\right] = Z(1-Z)(1-2Z).$$

Hence,

$$d_W(\bar{X}_n, \mathcal{N}(0,1)) \le 2(\sqrt{2}+1)\mathbb{E}\left[\frac{1-2Z+2Z^2}{\sqrt{Z(1-Z)}}\right] n^{-1/2}.$$

It is a classical quantitative theorem in the theoretical Stein's method, but in some formulations, one supposes the random variables to have unit variance or normalization (Reinert, 1998; Chen et al., 2010, corollary 4.2.). There is quite a few papers investigating the lower constant before the absolute third moments in the upper bound for the Kolmogorov distance (more difficult to obtain usually because the solution of Stein's equations has less bounded derivatives) and various other probability distances (see for example Tyurin (2009) and the references therein). One can consider the L^1-norm of the difference of cumulative function of $\bar{X}_n$ and $\mathcal{N}(0,1)$ (Goldstein et al., 2010). Recently, with the same ambition to find general bounds for functional of independent random variables, Privault and Serafin (2018) derives a bound in Wasserstein-1 distance for sums of identically distributed independent random variables with a constant 2 which is lower than ours. We note that Dung (2018) uses the same gradient, but the constant here is better than 4 in his.

Hélène impossible fourth

Actually the integration by parts with carré du champ operator has been proved powerful to bypass combinatorial difficulties with Malliavin derivatives (Döbler et al., 2018) when deriving fourth moment limit theorems. They are not equivalent in our discrete settings (see also Döbler et al. (2018))

2.6.3 Abstract bounds for U-statistics

The chaos decomposition has a natural interpretation as a decomposition in terms of degenerate U-statistics.

Definition 2.6.5 (U-statistic (Hoeffding, 1948))**.** Let a family of measurable functions $h_I : E_I \to \mathbb{R}$. A U-statistic of degree (or order) p is defined for any $n \ge p$ by:

$$U = \sum_{I\in(A,p)} h_I(X_I) = \sum_{I\in(A,p)} W_I.$$

Definition 2.6.6 (Degenerate U-statistic)**.** A degenerate U-statistic of order $p > 1$ is a U-statistic of order p such that $\mathbb{E}\left[h_I(X_I^{\{a\}}, x_a)|Z\right] = 0$, for all $a \in A$ and $x_a \in E_a$.

The space of degenerate U-statistics is exactly $\mathfrak{C}_p$. Since we consider functionals given Z here-

after, h_I may be $\sigma(Z)$-measurable as well.
A convenient assumption in the proofs of quantitative limit theorems is the diffusiveness of the Markov generator at hand L, i.e. the associated carré du champ Γ_L satisfies for (F, G) in a dense algebra of $\operatorname{Dom} L$:

$$\Gamma_L(\phi(F), G) = \phi'(F)\Gamma_L(F, G).$$

Due to the discreteness of the Malliavin structure, the operator L is not diffusive, but it is close to. We devise the following pseudo chain rule.

Lemma 2.6.7 — First pseudo chain rule. *Let $\psi \in C^1(\mathbb{R}, \mathbb{R})$. Let $G \in \mathcal{A}$ and $F \in L^2(E_A)$ such that $\psi(F) \in \mathcal{A}$, then:*

$$\Gamma(\psi(F), G) = \frac{1}{2}\sum_{a\in A} \psi'(F)\mathbb{E}\left[(\Delta^{\{a\}'}F)(\Delta^{\{a\}'}G)\,|\,X, Z\right] + R_\psi(F, G), \tag{2.6.5}$$

where:

$$|R_\psi(F, G)| \leq \frac{\|\psi''\|_\infty}{4}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}'}G|(\Delta^{\{a\}'}F)^2\,|\,X, Z\right].$$

Proof of lemma 2.6.7. We write the Taylor expansion of ψ, and:

$$\begin{aligned}\mathbb{E}\left[(\psi(F^{\{a\}'}) - \psi(F))(G^{\{a\}'} - G)\,|\,X, Z\right] &= \mathbb{E}\left[\psi'(F)(\Delta^{\{a\}'}F)(\Delta^{\{a\}'}G)\,|\,X, Z\right] \\ &\quad + \mathbb{E}\left[(G^{\{a\}'} - G)r_\psi(F, F^{\{a\}'} - F)\,|\,X, Z\right].\end{aligned}$$

Then,

$$\begin{aligned}2\Gamma(\psi(F), G) = \psi'(F)&\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)(\Delta^{\{a\}'}G)\,|\,X, Z\right] \\ &+ \sum_{a\in A}\mathbb{E}\left[(G^{\{a\}'} - G)r_\psi(F, F^{\{a\}'} - F)\,|\,X, Z\right]\end{aligned}$$

where:

$$r_\psi(x, y) = \psi(x + y) - \psi(x) - \psi'(x)y = \int_0^y (y - s)\psi''(x + s)ds.$$

We note that r_ψ satisfies:

$$|r_\psi(x, y)| \leq \frac{\|\psi''\|_\infty}{2}y^2,$$

and we obtain the bound on the remainder. □

Theorem 2.6.8 — Bounds in 1-Wasserstein distance. *Assume that $F \in L^3(E_A)$, such that*

$\mathbb{E}[F|Z] = 0$ *and* $\mathbb{E}[F^2] = 1$*, then we get the bound:*

$$d_W(F, \mathcal{N}(0,1)) \leq \sqrt{\frac{2}{\pi}}\mathbb{E}|\Gamma(F, -\mathsf{L}^{-1}F) - 1| + \frac{1}{2}\sum_{a \in A}\mathbb{E}[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|(\Delta^{\{a\}'}F)^2]. \quad (2.6.6)$$

Moreover, if $F \in L^4(E_A)$*, then one has the further bound:*

$$d_W(F, \mathcal{N}(0,1)) \leq \sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}(\Gamma(F, \mathsf{L}^{-1}F))} + \frac{\sqrt{2}}{2}\sqrt{-\mathbb{E}[F\mathsf{L}F]}\sqrt{\sum_{a \in A}\mathbb{E}[|\Delta^{\{a\}'}F|^4]}. \quad (2.6.7)$$

Proof of theorem 2.6.8. We have:

$$\begin{aligned}
\mathbb{E}[L^\dagger f^\dagger F] &= \mathbb{E}[F(f^\dagger)'(F) - (f^\dagger)''(F)] \\
&= \mathbb{E}[\mathsf{L}\mathsf{L}^{-1}F(f^\dagger)'(F)] - \mathbb{E}[(f^\dagger)''(F)] \\
&= \mathbb{E}[\mathsf{L}^{-1}F\mathsf{L}((f^\dagger)'(F))] - \mathbb{E}[(f^\dagger)''(F)] \\
&= \mathbb{E}[\Gamma(\mathsf{L}^{-1}((f^\dagger)'(F)), -\mathsf{L}^{-1}F)] - \mathbb{E}[(f^\dagger)''(F)]
\end{aligned} \quad (2.6.8)$$

by integration by parts. We use lemma 2.6.7 and obtain that:

$$\mathbb{E}[\Gamma(\mathsf{L}((f^\dagger)'(F)), -\mathsf{L}^{-1}F)] \leq \mathbb{E}[(f^\dagger)''(F)\Gamma(F, -\mathsf{L}^{-1}F)] + \mathbb{E}[R_{(f^\dagger)^{(3)}}(F, -\mathsf{L}^{-1}F)].$$

Thus,

$$\mathbb{E}[L^\dagger f^\dagger F] \leq \sqrt{\frac{2}{\pi}}\mathbb{E}|\Gamma(F, -\mathsf{L}^{-1}F) - 1| + \frac{1}{2}\sum_{a \in A}\mathbb{E}[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|(\Delta^{\{a\}'}F)^2].$$

By Jensen's inequality for the first term and Cauchy-Schwarz inequality (for expectation of sum of random variables) for the second one, then by integration by parts, it yields:

$$\mathbb{E}[L^\dagger f^\dagger F] \leq \sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}(\Gamma(F, \mathsf{L}^{-1}F))} + \frac{1}{2}\sqrt{\sum_{a \in A}\mathbb{E}[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|^2]}\sqrt{\sum_{a \in A}\mathbb{E}[(\Delta^{\{a\}'}F)^4]},$$

and the proof is complete. □

Corollary 2.6.9 *If* $F = \sum_{p=1}^m F_p$ *is four times integrable functional where* $F_p \in \ker(\mathsf{L}+p\mathrm{Id})$*,* *then:*

$$d_W(F, \mathcal{N}(0,1)) \leq \sqrt{\frac{2}{\pi}} \sum_{p,q=1}^{m} \frac{1}{q} \sqrt{\operatorname{Var}\left[\Gamma(F_p, F_q)\right]}$$
$$+ \sqrt{2} \sum_{p=1}^{m} \frac{1}{p} \sqrt{\mathbb{E}[F_p^2]} \left\{ \sum_{p=1}^{m} p^{1/4} \left(\sum_{a \in A} \mathbb{E} \left| \Delta^{\{a\}'} F \right|^4 \right)^{1/4} \right\}^2 . \quad (2.6.9)$$

Proof of corollary 2.6.9. We use the decomposition of L^{-1} as to develop the first and second terms in (2.6.7). The final result is obtained after using Cauchy-Schwarz inequality. □

That is the starting point towards a partial fourth moment limit theorem.

Remark 2.6.10. With that approach, we cannot state a Berry-Esseen bound because of the lack of smoothness of the solutions of the Stein's equation for the Kolmogorov distance.

Now, we turn to bounds in Kolmogorov distance which are based on the same computations. We recall the properties of the Stein's equation for Kolmogorov bounds (Chen et al., 2010, lemma 2.3.).

Lemma 2.6.11 — Inequalities for Kolmogorov test functions. *Let $z \in \mathbb{R}$, the test functions f_z in the Stein equation for Kolmogorov distance is such that:*

$$\|f_z\|_\infty \leq \frac{\sqrt{2\pi}}{4}, \|f_z'\|_\infty \leq 1.$$

Moreover, for all $u, v, w \in \mathbb{R}$,

$$|(w+u)f_z(w+u) - (w+v)f_z(w+u)| \leq \left(|w| + \frac{\sqrt{2\pi}}{4} \right) (|u| + |v|), \quad (2.6.10)$$

and the following local estimate holds for every $x, h \in \mathbb{R}$:

$$|f_z(x+h) - f_z(x) - hf_z'(x)| \leq \frac{h^2}{2} \left(|x| + \frac{\sqrt{2\pi}}{4} \right) + h(\mathbb{1}_{\{[x,x+h)]\}}(z) - \mathbb{1}_{\{x+h,x)\}}(z))$$
$$= \frac{h^2}{2} \left(|x| + \frac{\sqrt{2\pi}}{4} \right) + |h|(\mathbb{1}_{\{[x,x+h)]\}}(z) + \mathbb{1}_{\{x+h,x)\}}(z)). \quad (2.6.11)$$

Proposition 2.6.12 — Kolmogorov bounds. *Let $F \in L^4(E_A)$, then one has the bound:*

$$d_{Kol}(F, \mathcal{N}(0,1)) \leq \mathbb{E}|\Gamma(F, -\mathsf{L}^{-1}F) - 1| + \frac{\sqrt{2\pi}}{16} \sum_{a \in A} \mathbb{E}\left[|\Delta^{\{a\}'} \mathsf{L}^{-1} F| (F - F^{\{a\}'})^2 \right]$$
$$+ \frac{1}{4} \sum_{a \in A} \mathbb{E}\left[|\Delta^{\{a\}'} \mathsf{L}^{-1} F| |F| (F - F^{\{a\}'})^2 \right] + \frac{1}{2} \sup_{z \in \mathbb{R}} \sum_{a \in A} \mathbb{E}\left[|\Delta^{\{a\}'} \mathsf{L}^{-1} F| \Delta^{\{a\}'} F \Delta^{\{a\}'} \mathbb{1}_{\{F > z\}} \right] . \quad (2.6.12)$$

Proof. Using (2.6.11) in the case $x = F$, and $h = -\Delta^{\{a\}'}F$, one sees that, for every $a \in A$,

$$\begin{aligned}|f_z(F^{\{a\}'}) - f_z(F) - (F^{\{a\}'} - F)f_z'(F)| &\le \frac{(F - F^{\{a\}'})^2}{2}\left(|F| + \frac{\sqrt{2\pi}}{4}\right)\\ &+ |F - F^{\{a\}'}|(\mathbb{1}_{\{[F,F^{\{a\}'})]\}}(z) - \mathbb{1}_{\{F^{\{a\}'},F)\}}(z))\\ &\le \frac{(F - F^{\{a\}'})^2}{2}\left(|F| + \frac{\sqrt{2\pi}}{4}\right)\\ &+ (F^{\{a\}'} - F)(\mathbb{1}_{\{F^{\{a\}'}>z\}} - \mathbb{1}_{\{F>z\}}),\end{aligned}$$

Hence, plugging that into the Stein's equation:

$$\begin{aligned}\mathbb{E}[L^\dagger f^\dagger F] &= \mathbb{E}[F(f^\dagger)'(F) - (f^\dagger)''(F)]\\ &= \mathbb{E}[Ff_z(F) - f_z'(F)].\\ &= \mathbb{E}[F\mathsf{L}\mathsf{L}^{-1}(f_z(F))] - \mathbb{E}[f_z'(F)]\\ &= \mathbb{E}[\Gamma(-\mathsf{L}^{-1}F, f_z(F))] - \mathbb{E}[f_z'(F)] \text{ using the carré du champ integration by parts.}\end{aligned}$$

The integration by parts allows us to express $\mathbb{E}[L^\dagger f^\dagger F]$ without worrying about integration by $\mathbb{E}'$, instead of writing $\mathsf{L} = -\delta D$. Once more, it proves useful as to simplifying the bounds.

$$\begin{aligned}2\mathbb{E}[Ff_z(F)] &= \mathbb{E}[\sum_{a\in A}(\mathsf{L}^{-1}F^{\{a\}'} - \mathsf{L}^{-1}F)[f_z(F) - f_z(F^{\{a\}'})]] \text{ because of (??)}\\ &\le \sum_{a\in A}\mathbb{E}\left[|\mathsf{L}^{-1}F - \mathsf{L}^{-1}F^{\{a\}'}F|(F - F^{\{a\}'})f_z'(F)\right]\\ &+ \mathbb{E}\left[|\mathsf{L}^{-1}F - \mathsf{L}^{-1}F^{\{a\}'}|\frac{(F - F^{\{a\}'})^2}{2}\left(|F| + \frac{\sqrt{2\pi}}{4}\right)\right]\\ &+ \mathbb{E}\left[|\mathsf{L}^{-1}F - \mathsf{L}^{-1}F^{\{a\}'}|\Delta^{\{a\}'}F(\mathbb{1}_{\{F>z\}} - \mathbb{1}_{\{F^{\{a\}'}>z\}})\right]\\ &\le 2\mathbb{E}\left[\Gamma(-\mathsf{L}^{-1}F, F)f_z'(F)\right] + \mathbb{E}\left[|\mathsf{L}^{-1}F - \mathsf{L}^{-1}F^{\{a\}'}|\left(\frac{|F|}{2} + \frac{\sqrt{2\pi}}{8}\right)(F - F^{\{a\}'})^2\right]\\ &+ \sup_{z\in\mathbb{R}}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|\Delta^{\{a\}'}F\Delta^{\{a\}'}\mathbb{1}_{\{F>z\}}\right].\end{aligned}$$

One also have the further bounds:

$$
\begin{aligned}
&d_{Kol}(F,\mathcal{N}(0,1)) \leq \mathbb{E}|\Gamma(F,\mathsf{L}^{-1}F)-1| \\
&+\frac{1}{2}\left(\frac{\sqrt{\pi}}{4}(-\mathbb{E}[F\mathsf{L}F)])^{1/2}+\frac{1}{2}\left(\mathbb{E}[F^4]\mathbb{E}\left[\left(\sum_{a\in A}(\Delta^{\{a\}'}\mathsf{L}^{-1}F)^2\right)^2\right]\right)^{1/4}\right)\left(\sum_{a\in A}\mathbb{E}[|F^{\{a\}'}-F|^4]\right)^{1/2} \\
&+\frac{1}{2}\sup_{z\in\mathbb{R}}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|\Delta^{\{a\}'}F\Delta^{\{a\}'}\mathbb{1}_{\{F>z\}}\right] \\
&\leq \mathbb{E}|\Gamma(F,L^{\circ-1}F)-1| \\
&+\left(\frac{\sqrt{\pi}}{2}(-\mathbb{E}[F\mathsf{L}F)])^{1/2}+\frac{\sqrt{2}}{2}\frac{1}{2}\left(\mathbb{E}[F^4]\mathbb{E}\left[\left(\sum_{a\in A}(\Delta^{\{a\}'}\mathsf{L}^{-1}F)^2\right)^2\right]\right)^{1/4}\right)\left(3\mathbb{E}[F^2\Gamma(F,F)]+\mathbb{E}[F^3\mathsf{L}F]\right)^{1/2} \\
&+\sup_{z\in\mathbb{R}}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|\Delta^{\{a\}'}F\Delta^{\{a\}'}\mathbb{1}_{\{F>z\}}\right]
\end{aligned}
$$

We use the following inequality to bound $\mathbb{E}[F^4]$ given that $\mathbb{E}[F^2]=1$.

$$
\begin{aligned}
\mathbb{E}[F^4] &= \mathbb{E}[F^2]^2+\mathrm{Var}[F^2] \\
&= 1+\mathbb{E}[\Gamma(F^2,-\mathsf{L}^{-1}F^2)]. \\
&\leq 1+\sum_{a\in A}\mathbb{E}[|D_aF^2|^2],
\end{aligned}
$$

using Poincaré's inequality. Then, we use Jensen's inequality to bound the other term.

$$
\mathbb{E}\left[\left(\sum_{a\in A}(\Delta^{\{a\}'}\mathsf{L}^{-1}F)^2\right)^2\right] \leq \mathbb{E}\left[\sum_{a\in A}(\Delta^{\{a\}'}\mathsf{L}^{-1}F)^4\right].
$$

□

Because of that, we focus exclusively on the Wasserstein-1 distance in the remainder of this chapter.

Because our chaos decomposition coincides with the one of Döbler et al. (2017) in the case where $(X_i)_{i\in\mathbb{N}}$ is a collection of independent random variables, we can use the theorem 1.3. of that paper, but we show it with slightly different constants. The use of Mehler's formula as in Zheng (2019) yields smaller constant, but cannot be generalized to functionals with finite chaotic decomposition contrarily to ours (see also Döbler (2023)).

Using the decomposition of the carré du champ operator, we have the following theorem from Döbler (2023). In the following we assume that it exists $p\in\mathbb{N}$ such that $F\in\oplus_{m=0}^{p}\mathfrak{C}_m$.

Theorem 2.6.13 — (Döbler, 2023). *If $F = \sum_{p=1}^m$ where $F_p \in \mathfrak{C}_p$, then:*

$$d_{KR}(F, \mathcal{N}(0,1)) \leq \sqrt{\frac{2}{\pi}} \sqrt{\operatorname{Var}\left[\sum_{p,q=1}^m \frac{1}{q}\Gamma(F_p, F_q)\right]} + \sqrt{2}\sum_{p=1}^m \frac{1}{p}\sqrt{\mathbb{E}[F_p^2]} \left\{ \sum_{p=1}^m p^{1/4} \left(\sum_{a\in A} \mathbb{E}\left|F_p(X_A) - F_p(X_{A\setminus\{a\}}, X_a')\right|^4 \right)^{1/4} \right\}^2 \tag{2.6.13}$$

or alternatively,

$$d_W(F, \mathcal{N}(0,1)) \leq \sqrt{\frac{2}{\pi}} \sqrt{\operatorname{Var}\left[\sum_{p,q=1}^m \frac{1}{q}\Gamma(F_p, F_q)\right]} + \sqrt{2}\sum_{p=1}^m \frac{1}{p}\sqrt{\mathbb{E}[F_p^2]} \left(\sum_{p=1}^m p^{1/4}([3\mathbb{E}[\frac{1}{p}F_p^2\Gamma(F_p,F_p)] - \mathbb{E}[F_p^4]])^{1/4} \right)^2. \tag{2.6.14}$$

Definition of maximal influence

Theorem 2.6.14 — Döbler (2023). *Under the above assumptions, we have the following general bounds:*

$$\begin{aligned} d_W(F, \mathcal{N}(0,1)) &\leq \sqrt{\frac{2}{\pi}} \frac{2p-1}{2p} \left(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + \kappa_p \mathbb{E}[F^2]\rho_A^2\right)^{1/2} \\ &\quad + \sqrt{2}\frac{2p-1}{p} \left(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + \kappa_p \mathbb{E}[F^2]\rho_A^2 + \kappa_p \mathbb{E}[F^2]\rho_A^2\right)^{1/2} \\ &\leq \left(\sqrt{\frac{2}{\pi}} + 2\sqrt{2}\right) \sqrt{\mathbb{E}[F^4] - 3\mathbb{E}[F^2]} + \left(\sqrt{\frac{2}{\pi}} + 4\sqrt{2}\right) \mathbb{E}[F^2]\sqrt{\kappa_p}\rho_A, \end{aligned} \tag{2.6.15}$$

where κ_p is a constant that depends only on p.

Corollary 2.6.15 *Let $F \in \mathfrak{C}_p$ such that $F \in L^4(E_A)$ with $\mathbb{E}[F] = 0$ and $\mathbb{E}[F^2] = 1$.*

$$d_{KR}(F, \mathcal{N}(0,1)) \leq \left(\sqrt{\frac{2}{\pi}} + 2\sqrt{2}\right) \sqrt{\mathbb{E}[F^4] - 3} + \left(\sqrt{\frac{2}{\pi}} + 4\sqrt{2}\right) \sqrt{\kappa_p}\rho_p \tag{2.6.16}$$

Theorem 2.6.16 — (Döbler et al., 2017). *Let $F \in L^4(E_A)$ be a degenerate U-statistic of* *order p (i.e. $F \in \mathfrak{C}_p$) of independent random variables. Then, it holds that:*

$$\begin{aligned} d_{KR}(F, \mathcal{N}(0,1)) &\leq \sqrt{\frac{2}{\pi}} \left(\mathbb{E}[F^4] - 3 + \kappa_p \rho_A^2\right)^{1/2} + \frac{2\sqrt{2}}{3} \left(2(\mathbb{E}[F^4] - 3) + 3\kappa_p \rho_n^2\right)^{1/2} \\ &\leq \left(\sqrt{\frac{2}{\pi}} + \frac{4}{3}\right) \sqrt{|\mathbb{E}[F^4] - 3|} + \sqrt{\rho_A} \left(\sqrt{\frac{2}{\pi}} + 2\sqrt{2}\right) \rho_A, \end{aligned} \tag{2.6.17}$$

where κ_p is a constant which only depends on p, and $\rho_n = \max_{1\leq i\leq n} \sum_{\substack{K\subset[n],|K|=p \\ K\ni i}} \sigma_K^2 = \max_{1\leq i\leq n} \sum_{\substack{K\subset[n],|K|=p \\ K\ni i}} \mathbb{E}[W_K^2]$. Moreover, we have the following Kolmogorov bounds.

$$\begin{aligned} d_{Kol}(F, \mathcal{N}(0,1)) &\leq \left(\mathbb{E}[F^4] - 3 + \kappa_p \rho_n^2\right)^{1/2} + 2\sqrt{2} \left(2(\mathbb{E}[F^4] - 3) + 3\kappa_p \rho_n^2\right)^{1/2} \\ &\leq \left(\sqrt{\frac{2}{\pi}} + \frac{4}{3}\right) \sqrt{|\mathbb{E}[F^4] - 3|} + \sqrt{\rho_p} \left(\sqrt{\frac{2}{\pi}} + 2\sqrt{2}\right) \rho_A. \end{aligned} \tag{2.6.18}$$

To correct

2.7 Partial fourth moment theorems

We adapt the proof of Azmoodeh et al. (2014), requiring a second pseudo chain rule that expresses the carré du champ operator as an approximation of a derivation operator in its two arguments.

Lemma 2.7.1 — Second pseudo chain rule. *Let φ, ψ be twice differentiable functions such that their second derivative is bounded Lipschitz-continuous. Assume that F a four times integrable functional such that $\varphi(F) \in \mathcal{A}$, $F \in \mathcal{A}$ and $\mathbb{E}[F|Z] = 0$, then one has:*

$$\begin{aligned} \Gamma(\varphi(F), \psi(F)) &= (\varphi'\psi')(F)\Gamma(F,F) \\ &\quad - \frac{1}{4}(\varphi''\psi' + \varphi'\psi'')(F) \sum_{a\in A} \mathbb{E}\left[(\Delta^{\{a\}'}F)^3 | X, Z\right] + \sum_{a\in A} R_a, \end{aligned} \tag{2.7.1}$$

with:

$$R_a = \frac{1}{2}\left(\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) | X, Z\right] - \varphi(F)\mathbb{E}\left[R^{(4)}_{a,\psi}(F) | X, Z\right] - \psi(F)\mathbb{E}\left[R^{(4)}_{a,\varphi}(F) | X, Z\right]\right)$$

and:

$$R^{(4)}_{a,\psi} \leq \frac{\|\psi^{(4)}\|_\infty}{24} \mathbb{E}\left[(\Delta^{\{a\}'}F)^4 | X, Z\right] \text{ for any } \psi \text{ fourth times differentiable.}$$

Proof of lemma 2.7.1. We have:

$$\begin{aligned} &2\Gamma(\varphi(F), \psi(F)) \\ &\quad = 2\varphi'(F)\psi'(F)\Gamma(F,F) - \frac{3}{6}(\varphi''\psi' + \varphi'\psi'')(F) \sum_{a\in A} \mathbb{E}\left[(\Delta^{\{a\}'}F)^3 | X, Z\right] \\ &\quad + \sum_{a\in A} \mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - \varphi(F)R^{(4)}_{a,\psi}(F) - \psi(F)R^{(4)}_{a,\varphi}(F) | X, Z\right], \end{aligned} \tag{2.7.2}$$

with:

$$R^{(4)}_{a,\phi} = \frac{1}{6}\mathbb{E}\left[\int_F^{F^{\{a\}'}} \phi^{(4)}(x)(x-F)^4 \, \mathrm{d}x | X, Z\right],$$

for ϕ a four times differentiable function. □

We focus on functionals in the p-th chaos for $p > 0$, as to obtain such kind of bound:

$$\mathrm{Var}[\Gamma(F,F)] \leq C(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) + \text{remainder}.$$

Lemma 2.7.2 *Let $G \in \oplus_{k=0}^q \mathfrak{C}_k$. Then for any $\eta \geq q$,*

$$\mathbb{E}[G(\mathsf{L}+\eta Id)^2 G] \leq \eta \mathbb{E}[G(\mathsf{L}+\eta Id)G] \leq c\mathbb{E}[G(\mathsf{L}+\eta Id)^2 G], \tag{2.7.3}$$

where

$$c = \frac{1}{\eta - q} \wedge 1.$$

Proof of lemma 2.7.2. Since $G \in \oplus_{k=0}^q \mathfrak{C}_k$, we write

$$G = \sum_{k=0}^q \pi_k(G) \text{ and } \mathsf{L}G = -\sum_{k=0}^q k\pi_k(G). \tag{2.7.4}$$

It follows that

$$\begin{aligned}\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})^2 G] &= \mathbb{E}[G\mathsf{L}(\mathsf{L}+\eta\mathrm{Id})G] + \eta\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})G] \\ &= \mathbb{E}[G\sum_{k=0}^q k(k-\eta)\pi_k(G)] + \eta\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})G].\end{aligned}$$

By orthogonality of the chaos,

$$\mathbb{E}[G\sum_{k=0}^q k(k-\eta)\pi_k(G)] = -\mathbb{E}[\sum_{k=0}^q k(\eta-k)\pi_k(G)^2] \leq 0,$$

and the inequality holds in view on the assumption on η. In the same vein,

$$\begin{aligned}\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})G] &= \sum_{k=0}^q (\eta-k)\mathbb{E}[\pi_k(G)^2] \\ &\leq c\sum_{k=0}^q (\eta-k)^2\mathbb{E}[\pi_k(G)^2] \\ &= c\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})^2].\end{aligned}$$

Thus, it yields the result. □

Lemma 2.7.3 *For $F \in \mathfrak{C}_p \cap L^4(E_A)$ and Q a polynomial of degree two and $a > 0$,*

$$\mathbb{E}[Q(F)(\mathsf{L} + ap\mathrm{Id})Q(F)] = p\mathbb{E}\left[aQ^2(F) - \frac{Q'(F)F}{3Q''(F)}\right] - \mathbb{E}[R_Q(F)], \tag{2.7.5}$$

where R_Q is a remainder term that depends on Q. For $Q = H_2 = X^2 - 1$ the second Hermite polynomial, the remainder reads off:

$$\mathbb{E}[R_Q] = \mathbb{E}[R_{H_2}] = \frac{1}{6}\mathbb{E}\left[\sum_{a\in A} |\Delta^{\{a\}'}F|^4\right]. \tag{2.7.6}$$

Proof of lemma 2.7.3. We first integrate by parts, then use the pseudo chain rule of lemma 2.7.1:

$$\begin{aligned}\mathbb{E}[Q(F)\mathsf{L}Q(F)] &= -\mathbb{E}[\Gamma(Q(F), Q(F))] \\ &= -\mathbb{E}[Q'(F)^2\Gamma(F,F)] \\ &+ \frac{1}{6}(Q^2)^{(3)}(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3 | X, Z\right] \\ &- \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[R^{(4)}_{a,Q^2}(F)|X,Z\right] - 2Q(F)\mathbb{E}\left[R^{(4)}_{a,Q}(F)|X,Z\right]\right].\end{aligned} \tag{2.7.7}$$

Since $Q^{(3)} = 0$, we have:

$$\begin{aligned}\mathbb{E}[Q(F)\mathsf{L}Q(F)] &= -\mathbb{E}\left[[Q'(F)^2\Gamma(F,F)\right] \\ &+ \frac{1}{6}\mathbb{E}\left[(Q^2)^{(3)}(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3|X,Z\right]\right] \\ &- \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[R^{(4)}_{a,Q^2}(F)|X,Z\right]\right].\end{aligned} \tag{2.7.8}$$

Moreover,

$$\left(\frac{Q'(F)^3}{3Q''(F)}\right)' = \frac{3Q'(F)Q''(F)^2}{3Q''(F)^2} = Q'(F)^2. \tag{2.7.9}$$

Subsequently, we use the pseudo chain rule of lemma 2.7.1 taking $\psi = \mathrm{Id}$ and $\varphi = \frac{Q'(\cdot)^3}{3Q''(\cdot)}$:

$$\begin{aligned}\mathbb{E}[Q'(F)^2\Gamma(F,F)] &= \mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] \\ &+ \frac{1}{4}\mathbb{E}\left[(\varphi''\psi' + \varphi'\psi'')(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3|X,Z\right]\right] \\ &- \sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F)|X,Z\right] - \varphi(F)\mathbb{E}\left[R^{(4)}_{a,\psi}(F)|X,Z\right]\right] \\ &- \mathbb{E}\left[F\mathbb{E}\left[R^{(4)}_{a,\varphi}(F)|X,Z\right]\right] \\ &= \mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] + \frac{1}{4}\mathbb{E}\left[(Q'(\cdot)^2)'(F)\sum_{a\in A}(\Delta^{\{a\}'}F)^3\right] \\ &- \sum_{a\in A}\frac{1}{2}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - FR^{(4)}_{a,\varphi}(F)\right].\end{aligned} \tag{2.7.10}$$

Finally,

$$\begin{aligned}\mathbb{E}[Q(F)\mathsf{L}Q(F)] &= -\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] \\ &\quad + \mathbb{E}\left[\left(\frac{1}{4}(Q'(\cdot)^2)'(F) - \frac{1}{12}(Q^2)^{(3)}(F)\right)\sum_{a\in A}(\Delta^{\{a\}'}F)^3\right] \\ &\quad + \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - R^{(4)}_{a,Q^2}(F) - FR^{(4)}_{a,\varphi}(F)\right] \\ &= -\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] \\ &\quad + \mathbb{E}\left[\left(\frac{1}{4}(Q'(\cdot)^2)'(F) - \frac{1}{12}(Q^2)^{(3)}(F)\right)\sum_{a\in A}(\Delta^{\{a\}'}F)^3\right] \\ &\quad + \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - R^{(4)}_{a,Q^2}(F)\right].\end{aligned} \tag{2.7.11}$$

Because $F \in \mathfrak{C}_p$, we have: $-\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] = \mathbb{E}[\frac{Q'(F)^3}{3Q''(F)}\mathsf{L}F] = -p\mathbb{E}\left[\frac{Q'(F)^3}{3Q''(F)}F\right]$. For $Q = H_2 = X^2 - 1$ the second Hermite polynomial,

$$\frac{Q'(F)^3}{3Q''(F)} = \frac{4}{3}X^3,$$

so $\left(\frac{Q'(\cdot)^3}{3Q''(\cdot)}\cdot\right)^{(4)} = 32$ and $(Q^2)^{(4)} = 24$. Thus,

$$\sum_{a\in A}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - R^{(4)}_{a,Q^2}(F)\right] = \frac{(32-24)}{24}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}'}F|^4\right]. \tag{2.7.12}$$

Since $(Q'(\cdot)^2)'(F) = 8F$, and $(Q^2)^{(3)}(F) = 24F$, the result follows. □

The assumption under which a fourth moment theorem holds, is that $F \in \mathfrak{C}_p$ is a chaos eigenfunction with respect to the Markov generator L i.e.:

$$F^2 \in \oplus_{k=0}^{2p}\mathfrak{C}_k. \tag{EGF}$$

It is analog to the one in Ledoux et al. (2012); Azmoodeh et al. (2014). We show that it holds for an important class of U-statistics, homogeneous sums. We shall use the notation (A, p) that stands for the set of p-tuples of distinct elements of A.

Example 2.7.4 (Conditionally independent homogeneous sums). Let $p > 0$. If there exists $(a_I)_{I\subset A} \in \mathbb{R}^{\mathcal{P}(A)}$ such that

$$W = \sum_{k=1}^{p}\sum_{I\in(A,k)} a_I \prod_{i\in I} X_i, \tag{2.7.13}$$

then

1. W is square-integrable homogeneous sum of order p if X_i are $2p$-integrable. In that case, $W \in \mathcal{S}$.

2.
$$\mathbb{E}\left[W|Z\right] = \sum_{k=1}^{p} \sum_{I \in (A,k)} a_I \prod_{i \in I} \mathbb{E}\left[X_i|Z\right]$$

is a homogeneous sum of random variables $\hat{X}_i = \mathbb{E}\left[X_i|Z\right]$ for $i \in I$ with $I \in (A,k)$ for $k \leq p$.

Remark that $(a_I)_{I \subset A}$ may be a sequence of random variables, in which case there exists a family of functions $(g_I)_{I \subset A}$ such that $a_I = g_I(Z)$.

> **Lemma 2.7.5** *Let W a homogeneous sums of conditionally independent random variables given Z. Then* (EGF) *holds.*

Proof of lemma 2.7.5. Let us denote by W_I the component of F in (2.7.13) proportional to $\prod_{\alpha \in I} X_\alpha$. We want to prove that there exist $G_1, \ldots, G_{2p}$ with $G_i \in \mathfrak{C}_i \cup \{0\}$ such that $W_I W_J = \sum_{i=1}^{2p} G_i$. Note that if $I \cap J = \emptyset$, and $a \in I$, then a is not in J and vice versa. Therefore, $W_I W_J \in \mathfrak{C}_{|I|+|J|}$. In general,

$$\begin{aligned}
W_I W_J &\propto \prod_{\alpha \in I} Y_\alpha \prod_{\beta \in J} Y_\beta \\
&= \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in I \cap J} Y_\delta^2 \\
&= \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in I \cap J} \left(Y_\delta^2 - \mathbb{E}\left[Y_\delta^2|Z\right] + \mathbb{E}\left[Y_\delta^2|Z\right]\right) \\
&= \sum_{K \subset I \cap J} \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} \left(Y_\delta^2 - \mathbb{E}\left[Y_\delta^2|Z\right]\right) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}\left[Y_\delta^2|Z\right].
\end{aligned}$$

For $a \in A$:

$$\begin{aligned}
&\mathbb{E}\left[\prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} \left(Y_\delta^2 - \mathbb{E}\left[Y_\delta^2|Z\right]\right) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}\left[Y_\delta^2|Z\right] | \mathcal{G}_a^Z\right] \\
&\qquad = \begin{cases} 0 \text{ if } a \in K \cup ((I \setminus J) \cup (J \setminus I)) \\ \displaystyle\prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} \left(Y_\delta^2 - \mathbb{E}\left[Y_\delta^2|Z\right]\right) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}\left[Y_\delta^2|Z\right] \text{ otherwise.} \end{cases}
\end{aligned}$$

Hence, we get

$$\prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} \left(Y_\delta^2 - \mathbb{E}\left[Y_\delta^2|Z\right]\right) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}\left[Y_\delta^2|Z\right] \in \mathfrak{C}_{|K \cup (I \setminus J) \cup (J \setminus I)|}$$

with $|K \cup ((I \setminus J) \cap (J \setminus I))| \leq |I \cup J| \leq 2p$. Thus, (EGF) holds. □

Proposition 2.7.6 *For $F \in \mathfrak{C}_p \cap L^2(E_A)$ such that $\mathbb{E}[F^2] = 1$ and (EGF) holds, one has:*

$$\mathbb{E}[(\Gamma(F,F) - p)^2] \leq \frac{p^2}{3}(\mathbb{E}[F^4] - 3) + \frac{p}{12}\mathbb{E}\left[\sum_{a\in A} |\Delta^{\{a\}'} F|^4\right]. \tag{2.7.14}$$

Proof of proposition 2.7.6. By the very definition of Γ, one has:

$$\begin{aligned}\Gamma(F,F) - p &= \frac{1}{2}\mathsf{L}(F^2) - F\mathsf{L}F - p = \frac{1}{2}\mathsf{L}(F^2) + pF^2 - p \text{ for } F \in \mathfrak{C}_p\\ &= \frac{1}{2}(\mathsf{L} + 2p\mathrm{Id})(F^2 - 1).\end{aligned}$$

It follows that:

$$\mathbb{E}[(\Gamma(F,F) - p)^2] = \frac{1}{4}\mathbb{E}[\big((\mathsf{L} + 2p\mathrm{Id})(F^2 - 1)\big)^2].$$

Since L is a self-adjoint operator, this yields:

$$\mathbb{E}[(\Gamma(F,F) - p)^2] = \frac{1}{4}\mathbb{E}[H_2(F)(\mathsf{L} + 2p\mathrm{Id})^2 H_2(F)].$$

As (EGF) holds, we are in position to apply lemma 2.7.2 with $q = 2p$ and $\eta = 2p$:

$$\mathbb{E}[(\Gamma(F,F) - p)^2] \leq \frac{p}{2}\mathbb{E}[H_2(F)(\mathsf{L} + 2p\mathrm{Id})H_2(F)]. \tag{2.7.15}$$

According to lemma 2.7.3, with $a = 2$,

$$\begin{aligned}\frac{p}{2}\mathbb{E}[H_2(F)(\mathsf{L} + 2p\mathrm{Id})H_2(F)] &= \frac{p^2}{2}\mathbb{E}\left[2(F^2-1)^2 - \frac{4}{3}F^4\right] + \frac{p}{2}\mathbb{E}[R_{H_2}(F)]\\ &= \frac{p^2}{6}\mathbb{E}\left[6(F^2-1)^2 - 4F^4\right] + \frac{p}{2}\mathbb{E}[R_{H_2}]\\ &= \frac{p^2}{3}\mathbb{E}[F^4 - 6F^2 + 3] + \frac{p}{2}\mathbb{E}[R_{H_2}].\end{aligned}$$

Thus, it yields

$$\mathbb{E}[(\Gamma(F,F) - p)^2] \leq \frac{p^2}{3}\mathbb{E}[F^4 - 6F^2 + 3] + \frac{p}{2}\mathbb{E}[R_{H_2}], \tag{2.7.16}$$

and the proof is complete, using again lemma 2.7.3. □

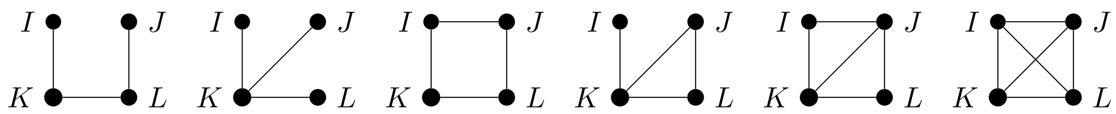
The remainder is also a fourth moment term.
Many papers are devoted to find the optimal conditions for the asymptotic normality of U-statistics. The criterion established in De Jong (1990) is related to the fourth moment phenomenon. The extra assumption is a negligibility condition also known as the Lindeberg-Feller condition. Fix A_m a finite subset of cardinal m such that $F = F(X_{A_m})$ and $\mathbb{E}[F^2] = 1$, that means:

$$\rho_{A_m}^2 = \max_{i\in A_m} \sum_{I\ni i,\, I\subseteq A_m,\, |I|=p} \mathbb{E}[W_I^2] \xrightarrow{m\to+\infty} 0. \tag{2.7.17}$$

In some papers Döbler et al. (2017), the term ρ_{A_m} is called maximal influence of the random variables on the total variance of the degenerate U-statistics F. In the following, we shall denote it by ρ. The condition (2.7.17) is not necessary for asymptotic normality to hold, but there exist counterexamples for which the sequence of fourth cumulants of functionals of independent Rademacher random variables converges to 0 while (2.7.17) does not hold (see Döbler et al. (2019)).
We show that the quantity is related to the remainder above.

Definition 2.7.7 (Connectedness of subsets)**.** The r-tuple $(I_1, \ldots, I_r)$ subsets of A is connected if the intersection graph of $\{I_1, \ldots, I_r\}$ is connected, i.e. the graph G with vertex set $\{I_1, \ldots, I_r\}$ and edge set $E(G) = \{\{I_i, I_j\} | \, i \neq j, I_i \cap I_j \neq \emptyset\}$ is connected.

In the case where $r = 4$, there are exactly six simple connected graphs with only four vertices (up to isomorphisms). There are listed in the figure below. The edges mean that the intersection between two vertices is non empty.



In that order, they are respectively: linear (4-path), 4-star, square, kite (tadpole), diag (diamond), and 4-complete graph K_4.

Lemma 2.7.8 *If $F \in \mathfrak{C}_p \cap L^4(E_A)$, then:*

$$\sum_{a \in A} \mathbb{E}[|\Delta^{\{a\}'} F|^4] \leq 16p \sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|. \tag{2.7.18}$$

Moreover, assuming the hypercontractivity condition, i.e.

$$\sup_{J \in (A,p)} \frac{\mathbb{E}[W_J^4]}{\mathbb{E}[W_J^2]^2} < +\infty, \tag{HC}$$

there exists a constant c_p that depends only on p such that:

$$\sum_{a \in A} \mathbb{E}[|\Delta^{\{a\}'} F|^4] \leq c_p \rho^2. \tag{2.7.19}$$

Proof of lemma 2.7.8. Because $(a+b)^4 \leq 8(a^4 + b^4)$, one has:

$$\begin{aligned}
\sum_{a \in A} \mathbb{E}\left|\Delta^{\{a\}'} F\right|^4 &\leq 8 \sum_{a \in A} \mathbb{E}\left[\left(\sum_{I \ni a, |I| \leq p} W_I^{\{a\}'}\right)^4 + \left(\sum_{I \ni a, |I| \leq p} W_I\right)^4\right] \\
&= 16 \sum_{a \in A} \mathbb{E}\left[\left(\sum_{I \ni a, |I| \leq p} W_I\right)^4\right] \\
&\leq 16 \sum_{I \cap J \cap K \cap L \neq \emptyset} |I \cap J \cap K \cap L| \mathbb{E}[W_I W_J W_K W_L] \\
&\leq 16p \sum_{I \cap J \cap K \cap L \neq \emptyset} |\mathbb{E}[W_I W_J W_K W_L]| \\
&\leq 16p \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|
\end{aligned}$$

Then, we bound it by the maximal influence, using the generalized Hölder inequality:

$$\begin{aligned}|\mathbb{E}[W_I W_J W_K W_L]| &\leq \left(\mathbb{E}[W_I^4]\mathbb{E}[W_J^4]\mathbb{E}[W_K^4]\mathbb{E}[W_L^4]\right)^{1/4} \\ &\leq \max_{J\in A, |J|=p} \frac{\mathbb{E}[W_J^4]}{\mathbb{E}[W_J^2]^2} \left(\mathbb{E}[W_I^2]^2\mathbb{E}[W_J^2]^2\mathbb{E}[W_K^2]^2\mathbb{E}[W_L^2]^2\right)^{1/4}\end{aligned}$$

with $\sigma_I^2 = \mathbb{E}[W_I^2]$. Then the proposition 2.9 of Döbler et al. (2017) can be extended for functionals of conditionally independent random variables and implies that:

$$\sum_{I\cap J\cap K\cap L\neq\emptyset} \sigma_I\sigma_J\sigma_K\sigma_L \leq C_p\rho^2,$$

where the finite constant C_p only depends on p. It yields the existence of $c_p > 0$ such that the inequality (2.7.19) holds true. □

We are now in position to state a partial fourth moment limit theorem.

Theorem 2.7.9 — Quantitative De Jong's limit theorem I. *Let $F \in L^4(E_A)$ a degenerate U-statistics of order p of conditionally independent random variables such that $\mathbb{E}[F|Z] = 0$ and $\mathbb{E}[F^2] = 1$. If we suppose the hypercontractivity condition* (HC) *and the assumption* (EGF)*, then one has the bound:*

$$d_W(F, \mathcal{N}(0,1)) \leq \sqrt{\frac{2}{3\pi}}\sqrt{\mathbb{E}[F^4] - 3} + \tilde{C}_p\rho, \tag{2.7.20}$$

with $\tilde{C}_p$ a positive constant that only depends on p.

Proof. By corollary 2.6.9,

$$d_W(F, \mathcal{N}(0,1)) \leq \sqrt{\frac{2}{\pi}\frac{1}{p}}\sqrt{\mathrm{Var}\left[\Gamma(F,F)\right]} + \sqrt{2}\sqrt{\mathbb{E}[F^2]}\left(\sum_{a\in A}\mathbb{E}\left[\left|\Delta^{\{a\}'}F\right|^4\right]\right)^{1/2}.$$

The combination of (2.7.14) and lemma 2.7.8 yields the final upper bound. □

The upper bound of the remainder expressed in terms of maximal influence is not used in the subsequent applications, so we drop the (HC) condition.
A related result to the fourth moment phenomenon appears in De Jong (1996) in the particular case that is the topic of the next chapter. We prove the associated quantitative statement for functionals of conditionally independent random variables. We prepare the proof with the following proposition.

Proposition 2.7.10 *If $F = \sum_{p=1}^m F_p$ where $F_p = \sum_{|I|=p} W_I \in \mathfrak{C}_p$, assuming there exists $C \in \mathbb{R}^+$ such that for all $I, J \subset A$, and $a \in A$, that*

$$\frac{\mathbb{E}[W_I W_J | \mathcal{G}^a]}{W_{I\setminus\{a\}}W_{J\setminus\{a\}}} < C\,\mathbb{P}\text{-a.s.}, \tag{H1}$$

then for $p \neq q$:

$$\sqrt{\mathrm{Var}\left[\Gamma(F_p, F_q)\right]} \lesssim \sqrt{\sum_{(I,J,K,L)\text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|}, \tag{2.7.21}$$

for I, J, K, L sets of size less than $\max(p, q)$.

Proof of proposition 2.7.10. The carré du champ reads for $p \neq q$:

$$\begin{aligned}\Gamma(F_p, F_q) &= \Gamma(\sum_{|I|=p} W_I, \sum_{|J|=q} W_J)\\ &= \sum_{|I|,|J|=p,q} \Gamma(W_I, W_J)\end{aligned}$$

Hence,

$$\begin{aligned}2\Gamma(F_p, F_q) &= \sum_{|I|,|J|=p,q} (\mathsf{L}(W_I W_J) + (p+q) W_I W_J)\\ &= \sum_{|I|,|J|=p,q} \left((p+q) W_I W_J - \sum_{a\in A} D_a(W_I W_J)\right)\\ &= \sum_{|I|,|J|=p,q} \left((p+q) W_I W_J - \sum_{a\in I\cup J} D_a(W_I W_J)\right)\\ &= (p+q) \sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}} W_I W_J + \sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}} (|I| + |J| - |I\cup J|) W_I W_J\\ &+ \sum_{a\in I\cup J} \mathbb{E}\left[W_I W_J | \mathcal{G}_a\right].\end{aligned}$$

Because of the spectral decomposition, $\mathbb{E}\left[W_I | \mathcal{G}_a\right] = 0$ for $a \in I$. Let J such that $a \notin J$, then $\mathbb{E}\left[W_I W_J | \mathcal{G}_a\right] = W_J \mathbb{E}\left[W_I | \mathcal{G}_a\right] = 0$.

$$2\Gamma(F_p, F_q) = (p+q) \sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}} W_I W_J + \sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}} \sum_{a\in I\cap J} \left(W_I W_J + \mathbb{E}\left[W_I W_J | \mathcal{G}_a\right]\right).$$

Then for $p \neq q$, using the convexity of $x \longmapsto x^2$,

$$\begin{aligned}\mathrm{Var}(\Gamma(F_p, F_q)) \leq \frac{1}{2} \mathrm{Var}&\left[(p+q) \sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}} W_I W_J\right]\\ &+ \frac{1}{2} \mathrm{Var}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}} \sum_{a\in I\cap J} \left(W_I W_J + \mathbb{E}\left[W_I W_J | \mathcal{G}_a\right]\right)\right]\end{aligned}$$

$$\mathrm{Var}(\Gamma(F_p,F_q)) \leq \frac{1}{2}\mathbb{E}\left[\left((p+q)\sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}} W_I W_J\right)^2\right] + \frac{1}{2}\,\mathrm{Var}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{a\in I\cap J}\left(W_IW_J+\mathbb{E}\left[W_IW_J\,|\,\mathcal{G}_a\right]\right)\right]$$

$$\begin{aligned}2\,\mathrm{Var}(\Gamma(F_p,F_q)) \leq{} & \sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L=\emptyset}} \mathbb{E}[W_IW_JW_KW_L] \\ & + \mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} W_IW_JW_KW_L\right] \\ & + \mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} W_IW_J\mathbb{E}\left[W_KW_L\,|\,\mathcal{G}_b\right]\right] \\ & + \mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} \mathbb{E}\left[W_IW_J\,|\,\mathcal{G}_a\right]W_KW_L\right] \\ & + \mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} \mathbb{E}\left[W_IW_J\,|\,\mathcal{G}_a\right]\mathbb{E}\left[W_KW_L\,|\,\mathcal{G}_b\right]\right].\end{aligned}$$

We shall write

$$|C_{I,J,a}| = \left|\frac{\mathbb{E}\left[W_IW_J\,|\,\mathcal{G}_a\right]}{W_{I\setminus\{a\}}W_{J\setminus\{a\}}}\right| \text{ for all } I,J,a$$

with the convention $W_\emptyset = 1$.
Let us deal with each term one by one:

- If $I\cap J=\emptyset$, $K\cap L=\emptyset$, and if there is more than 2 other pairs with null intersection, the contribution of the term is 0, hence the first term is non-zero if (I,J,K,L) is connected, then:
$$\sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L=\emptyset}} \mathbb{E}[W_IW_JW_KW_L] \leq \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_IW_JW_KW_L]|.$$
- The second term consists of the sums of product of factors indexed by connected sets since there are at least two pairs that have non-null intersection. Since $p\neq q$, $\mathbb{E}\left[W_IW_J\,|\,Z\right]=0$ for $|I|=p$ and $|J|=q$, so if the terms are non-zero, W_IW_J and W_KW_L are not conditionally independent.

- For the third term, using self-adjointness, the terms are non-zero if $b \in I \cap J$, hence it is equivalent to:

$$|C_{I,J,a}\mathbb{E}[W_{I\setminus\{b\}}W_{J\setminus\{a\}}W_KW_L]| = |C_{I,J,a}||\mathbb{E}[W_{I\setminus\{b\}}W_{J\setminus\{a\}}W_KW_L]|.$$

  If b is the unique element that lies in the intersection, the contribution is 0, otherwise I, J, K, L are connected or the contribution is

$$\mathbb{E}\left[W_IW_J \,|\, Z\right]\mathbb{E}\left[W_KW_L \,|\, Z\right] = 0$$

  because $|I| \neq |J|$.

- For the last term, it is the same argument.

Then, there exists a constant C independent of others such that

$$\mathrm{Var}(\Gamma(F_p, F_q)) \leq (1 + m^2 + 2Cm^2 + C^2m^2) \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_IW_JW_KW_L]|.$$

Remark 2.7.11. It si clear that the constant C_m depends on the distributions of the random variables, hence it is not a universal constant as found in Döbler et al. (2017); Döbler (2023). It is a generalization of the main result for normal approximation in Bhattacharya et al. (2022) to homogeneous sums. In the case where the random variables are identically distributed, the conditions for the result are convenient. However, in the case of homogeneous sums, the (H1) condition additionally requires that the random variables are square-integrable. Moreover, the constant depends on the order of the U-statistic which can detrimental to the convergence rate for U-statistics with non-finite Hoeffding decomposition.

□

In Privault and Serafin (2022), Privault and Serafin proves a partial fourth moment theorem for F a functional of independent random variables sum of element in the first and second chaos of their own Malliavin structure. To that end, we devise another strategy which is to reexpress the remainder in the partial fourth moment theorem as a fourth order term.

Theorem 2.7.12 — Quantitative De Jong's theorem II. *If $F = \sum_{p=1}^{m} F_p$ where $F_p \in \mathfrak{C}_p$ and let us assume:*

- *F_p are chaos eigenfunctions* (EGF)*;*
- *the condition* (H1)*;*
- $$\kappa = \sup_{I,J \subset A} \frac{\mathbb{E}[W_I^2]\mathbb{E}[W_J^2]}{\mathbb{E}[W_I^2W_J^2]} < \infty \tag{H2}$$

  is independent of A.

Then:

$$d_W(F, \mathcal{N}(0,1)) \leq C_m \sqrt{\sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_IW_JW_KW_L]|}, \tag{2.7.22}$$

where the constant C_m grows quadratically with m, independent of all others.

Proof of theorem 2.7.12. Let us prove the upper bound of Var $[\Gamma(F_p, F_p)]$ by bounding the fourth cumulant:

$$\begin{aligned}\mathbb{E}[F_p^4] = 3 \sum_{\substack{I,J,K,L\in(A,p)\\(I\cup J)\cap(K\cup L)=\emptyset}} \mathbb{E}[W_I W_J]\mathbb{E}[W_K W_L] + \sum_{\substack{I,J,K,L\in(A,p)\\I,J,K,L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] \\ = 3 \sum_{I,J\in(A,p)} \mathbb{E}[W_I^2]\mathbb{E}[W_J^2] - 3 \sum_{I\cap J\neq\emptyset\neq} \mathbb{E}[W_I^2]\mathbb{E}[W_J^2] \\ + \sum_{\substack{I,J,K,L\in(A,p)\\I,J,K,L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] \\ = 3\mathbb{E}[F_p^2]^2 + \sum_{\substack{I,J,K,L\in(A,p)\\I,J,K,L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] - 3 \sum_{\substack{I\cap J\neq\emptyset\\I\neq J}} \mathbb{E}[W_I^2]\mathbb{E}[W_J^2]\end{aligned}$$

Then, one has:

$$|\mathbb{E}[F_p^4] - 3|\mathbb{E}[F_p^2]^2| \leq (1+3\kappa) \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|. \tag{2.7.23}$$

□

The assumptions may seem cumbersome, but as shown in lemma 2.7.5 concerning (EGF), they are valid for homogeneous sums.

2.8 Comments

The Malliavin structure at hand is the extension of the one for independent random variables. As all the strategies for tackling the problem of approximation of independent random variables share the same use of underlying properties of the random variables such as the existence of exchangeable pairs, they complete each other in the range of applications across the literature. In the next chapter, we consider one of the applications in motif estimation in random hypergraphs.

Chapter 3

Motif estimation

Graphs are used in a lot of areas such as social network analysis, bio informatics, and telecommunication networks. Networks are a versatile mathematical tool for representing the structure of complex systems and have been the subjects of large volume of work Knoke and Yang (2019); Borgatti et al. (2009); Newman et al. (2006). Graphs that represent network can have directed edges, weighted edges or vertices, have several kinds of vertices like in bipartite networks. Our work focus on non-directed graph.

But some of those graphs are really huge in size, and learning about the structure of those graphs leads to lots of challenging questions. Indeed, we cannot store the whole information of some graphs in memory of a computer. For instance, in October 2012, Facebook reported to have 1 billions users. Using 8 bytes for user ID, 100 friends per user, storing the raw edges will take 1 Billions $\times 100 \times 8$ bytes = 800 GB. Today, the size of the Facebook graph is not likely to be processed offline. And that is only for the storage of basic information about users. Network analysis also involves going through users' media.

Some fields such as routing protocols require running simulation algorithms on those large graphs. Unfortunately, in addition to the occupation of space, the run time of algorithms on the whole set of nodes turns out to be too long as it scales in polynomial time with the number of nodes.

Then, instead of studying the wholeness of a given graph, we consider subgraphs or sampling of graph. In the latter case, the sampling process is often random. Even though some questions arise about obtaining good samples of graphs, we do not pursue in that approach.

Most often in network analysis, the observed data is a sample from a much larger parent network. The central statistical question in such studies is to estimate global features of the parent network with a control of the approximation made by sampling. Counting motifs (patterns of subgraphs) in a large network is a prominent statistical and computational problem. The overarching goal of this effort is to find sharp bounds of normal approximation in various subgraph counting problems, and extend the results to hypergraphs. We resort to fourth moment theorems.

3.1 Subhypergraph counting in random hypergraphs

Random graph models are various, but in the context of dense graphs, they all share a common property.

Introduce the notion of exchangeability small paragraph

The vertices of those random graphs are indistinguishable. That is to say, that they are informative of the structure of the random objects. It is related to the fact that the graph can be considered unlabelled in non-combinatorics applications.
The random hypergraphs are natural extensions of random graphs. A vast majority of the literature deals with the Erdös-Rényi model and its generalization. It is an example of exchangeable random hypergraphs.

Definition 3.1.1. A k-uniform exchangeable random hypergraph of vertex set $[n] = \{1, \ldots, n\}$ is defined by the set of $\{0,1\}$-valued random variables $(X_\alpha, \alpha \subset \binom{[n]}{k})$. One associates each realization of the random variables a hypergraph $([n], E)$ with $\alpha \in E$ if and only if $X_\alpha = 1$.

We can formulate a recipe for exchangeable random hypergraphs as done in (Austin, 2008, definition 2.8). Fix a sequence of ingredients

$$(\{*\}), (V, P_1), (\{0,1\}, P_2), (\{0,1\}, P_3), \ldots, (\{0,1\}, P_{k-1}), (\{0,1\}, P_k)$$

where $(P_k)_{k\in\mathbb{N}}$ a family of probability kernels such that for all $k \in \mathbb{N}$, P_k is a probability kernel from $\binom{V}{k-1}$ to $\binom{V}{k}$.

- Colour each vertex $s \in V$ by some $z_s \in \{0,1\}$ chosen independently according to $P_1(z_\emptyset, \cdot)$;
- Colour each edge $a = \{s,t\} \in \binom{V}{2}$ by some $z_a \in \{0,1\}$ chosen independently according to $P_2(z_\emptyset, z_s, z_t, \cdot)$;

$$\vdots$$

- Colour each $(k-1)$-hyperedge $u \in \binom{V}{k-1}$ by some $z_u \in \{0,1\}$ chosen independently according to $P_{k-1}(z_\emptyset, (z_s)_{s\in\binom{[u]}{1}}, \ldots, (z_v)_{v\in\binom{[u]}{k-2}}, \cdot)$;
- Colour each k-hyperedge $e \in \binom{V}{k}$ by some colour $z_e \in \{0,1\}$ chosen independently according to $P_k(z_\emptyset, (z_s)_{s\in\binom{[e]}{2}}, \ldots, (z_u)_{u\in\binom{[e]}{k-1}}, \cdot)$.

It consists in sampling first edges, then 3-hyperedges, and so on up to the rank k.

Remark 3.1.2. We have relative independence, as the the choices of hyperedges are independent relatively to the previous stages.

Example 3.1.3 (Erdös-Rényi random model)**.** The randomness intervenes at the level of edges. We colour each edge $a = \{s,t\} \in \binom{V}{2}$ by some $z_a \in Z_2 = \{0,1\}$ chosen independently according to $P_2(z_\emptyset, z_s, z_t, \cdot) \overset{d}{=} \mathcal{B}(p)$ for some $p \in [0,1]$.

Example 3.1.4 (Stochastic block model)**.** A stochastic block model corresponds to a model where there are communities, and each edge has a probability of belonging to the model according to the community of the vertices that the edge links. Likewise, the randomness intervenes at the level of the edges. Let a partition $V = C_1 \sqcup \ldots \sqcup C_q$. Let $(p_{i,j})_{i,j \in [\![1,q]\!]^2}$ a sequence of reals in $[0, 1]$. We can assign a community to each vertex s, let call it $c(s)$. Then:

- $P_1(z_\emptyset, c(s)) = \mathbb{1}_{\{s \in C_{c(s)}\}}$;
- $P_2(z_\emptyset, c(s), c(t), \cdot) \overset{d}{=} \mathcal{B}(p_{c(s),c(t)})$.

The natural extension of the Erdös-Rényi model denoted $\mathbb{G}^{(3)}(n, p_n)$ consists of having

$$P_3(z_\emptyset, z_{st}, z_{tu}, z_{us}) \overset{d}{=} \mathcal{B}(p_n),$$

i.e. we draw every triangle of the hypergraph with probability p_n. We also consider another random model based on the recipe. Let $(\mathbb{T}^{(3)}(n, q_n, p_n))_{n\in\mathbb{N}}$ the sequence of 3-uniform hypergraphs such that for $(s, t, u) \in V^3$:

- $$P_2(z_\emptyset, z_s, z_t) \overset{d}{=} \mathcal{B}(q_n);$$
- $$P_3(z_\emptyset, z_{st}, z_{tu}, z_{us}) \overset{d}{=} \mathcal{B}(p_n).$$

It differs from $\mathbb{G}^{(3)}(n, p_n)$ in many ways as pointed out by (Lovász, 2012, Example 23.11), but we note that $\mathbb{G}^{(3)}(n, p_n)$ and $\mathbb{T}^{(3)}(n, 1, p_n)$ have the same law. The case $q_n < 1$ has not been much studied in the literature. The functional identities in 2.5 can be applied to random hypergraphs in the same way as for random graphs (Janson et al., 2000, corollary 2.27).
One of the oldest problem of motif estimation is *subgraph counting* in random graphs or random hypergraphs. Central limit theorems for U-statistics suffice to prove conditions for asymptotic normality, but we can also derive convergence rates and a little more than convergence in law in general.
We also recall some definitions for sets and hypergraphs.

Definition 3.1.5. Let E a set (a vertex set for example), and denote E_r the set of all r-tuples with distinct indices. Thus, if E is finite, the cardinality of E_r is $\frac{|E|!}{(|E|-r)!}$.

Definition 3.1.6. The set $Aut(H)$ is the *automorphism group* of H a r-uniform hypergraph that is, the number of permutations σ on the vertex set $V(H)$ such that $(x_1, x_2, \ldots, x_r) \in E(H)$ if and only if $(\sigma(x_1), \sigma(x_2), \ldots, \sigma(x_r)) \in E(H)$.

Small subgraph counts can be used as summary statistics for large random graphs.

3.2 Poisson approximation

The Poisson approximation of subgraph count has been extensively studied (see Janson et al. (2000, example 6.26 p.161)) using Stein-Chen method and the well-known theorem for Poisson approximation that can be found for example in Janson et al. (2000, theorem 6.22. p.159) (see also the earlier reference Barbour et al. (1992)).

Theorem 3.2.1 — (Arratia, Goldstein, Gordon, et al., 1989; Barbour, Holst, and Janson, 1992). *Suppose that* $X = \sum_{\alpha \in A}$*, where* I_α *are the random indicator variables with* $p_\alpha = \mathbb{E}[I_\alpha]$*, and* $\lambda = \mathbb{E}[X] = \sum_\alpha p_\alpha$*. Suppose that there exists a family of random indicators* $J_{\beta\alpha}$*,* $\beta \in A \setminus \{\alpha\}$*, such that:*

$$J_{\beta\alpha} = I_\beta | I_\alpha = 1.$$

Then,

$$d_{TV}(\mathcal{L}(X), Po(\lambda)) = (1 \wedge \lambda^{-1}) \left(\sum_{\alpha \in \Lambda} p_\alpha^2 + \sum_{\alpha \in \Lambda} \sum_{\beta \neq \alpha} p_\alpha \mathbb{E}[|I_\beta - J_{\alpha\beta}|] \right). \tag{3.2.1}$$

One way to apply that theorem without explicit construction of the variables $J_{\beta\alpha}$ is via a dependency graph. In fact, if the family $\{I_\alpha\}$ has a dependency graph G, then there exists random variables $J_{\beta\alpha}$ with the right distribution such that $J_{\beta\alpha} = I_\beta$ when $\alpha\beta \notin E(G)$, so it suffices to consider $|J_{\beta\alpha} - I_\beta| \leq J_{\beta\alpha} + I_\beta$ together with the general relation

$$\pi_\alpha \mathbb{E}[J_{\beta\alpha}] = \pi_\alpha \mathbb{E}\left[I_\beta \,|\, I_\alpha = 1\right] = \mathbb{P}(I_\alpha = I_\beta = 1) = \mathbb{E}[I_\alpha I_\beta], \tag{3.2.2}$$

which yields:

$$\pi_\alpha \mathbb{E}|J_{\beta\alpha} - I_\beta| \leq \mathbb{E}[I_\alpha I_\beta] + \pi_{\alpha\beta}. \tag{3.2.3}$$

That leads to Janson et al. (2000, theorem 6.23. p.160), and even Janson et al. (2000, theorem 6.24. p.160).

Theorem 3.2.2 *Suppose that* $X = \sum_{\alpha \in A} I_\alpha$ *where the* I_α *are positively related random indicator variables. Then, with* $p_\alpha = \mathbb{E}[I_\alpha]$ *and* $\lambda = \mathbb{E}[X] = \sum_{\alpha \in A} p_\alpha$*.*

$$\begin{aligned} d_{TV}(X, Pois(\lambda)) &\leq 1 \wedge \lambda^{-1} \left(\operatorname{Var}(X) - \mathbb{E}[X] + 2 \sum_{\alpha \in A} p_\alpha^2 \right) \\ &\leq \frac{\operatorname{Var}[X]}{\mathbb{E}[X]} - 1 + 2 \max_{\alpha \in A} p_\alpha. \end{aligned} \tag{3.2.4}$$

Example 3.2.3 (Poisson approximation of the number of copies of a fixed pattern in Erdös-Rényi model)**.** Consider N_H the number of copies of a fixed r-uniform hypergraph H in $\mathbb{G}(n, p_n)$, and suppose H is strictly balanced (i.e. connected). Let us write the estimator as:

$$N_H = \sum_{\substack{G \in \binom{[n]}{2} \\ G \simeq H}} I_G,$$

where the random indicator variable is $I_G = \mathbb{1}_{\{G \subset \mathbb{G}(n, p_n)\}}$. We note that: $p_G = p_n^{e_G}$. The sums have $(1 + o(1))n^{v_H}/|Aut(H)|$ terms, each having expectation p^{e_H}. Thus,

$$\mathbb{E}[N_H] = n^{v_H} p_n^{e_H} / |Aut(H)| = \left(n p_n^{e_H/v_H}\right)^{v_H} = \left(n p_n^{d(H)}\right)^{v_H} \xrightarrow{n \to +\infty} \lambda.$$

Moreover, since H is strictly balanced, we have that for every proper subhypergraph H'

of H: $d(H') < d(H)$, then $\mathbb{E}[N_{H'}] \xrightarrow{n\to+\infty} +\infty$.

$$\begin{aligned}
\mathrm{Var}[N_H] &= \mathrm{Var}\left[\sum_{\substack{G\in\binom{[n]}{2}\\ G\simeq H}} I_G\right] \\
&= \sum_{\substack{G_1,G_2\in\binom{[n]}{2}\\ G_1,G_2\simeq H}} \mathrm{Cov}(I_{G_1}, I_{G_2}) \\
&= \sum_{E(G_1)\cap E(G_2)\neq\emptyset} (\mathbb{E}[I_{G_1}I_{G_2}] - \mathbb{E}[I_{G_1}]\mathbb{E}[I_{G_2}])
\end{aligned}$$

We use the computations in Nowicki and Wierman (1988, section 5). Let us denote for $d = 2, \ldots, e_H$, $f_d = |\{(A, B) : |E(A)\cap E(B)| = d, A, B \sim H\}|$ i.e. f_d is the number of pairs of subgraphs isomorphic to G with exactly d common edges. To compute f_d, decompose the set according to the number of common vertices, defining:

$$f_d(i) = |\{(A, B) : \ |E(A) \cap E(B)| = d,\ |V(A) \cap V(B)| = i,\ A, B \sim H\}|,$$

for $i = 3, 4, \ldots, v_H$ (since A and B share two or more common edges, they have at least 3 common vertices). Then,

$$f_d = \sum_{i=3}^{v_H} f_d(i).$$

For each $i = 3, \ldots, v_H$, we choose the i common vertices and the $v_H - i$ additional vertices in each A and B, so:

$$f_d(i) = \binom{n}{i, v_H - i, v_H - i} e_d(i),$$

where $e_d(i)$ denotes the number of pairs (A, B) which can be obtained on two fixed sets of vertices $V_1 = V(A)$ and $V_2 = V(B)$ such that: $|V_1 \cap V_2| = i$ and $|E(A) \cap E(B)| = d$. Since $e_d(i) =$, $i = 3, 4, \ldots, v_H$ is a sequence of constants independent of n,

$$\begin{aligned}
f_d &= \sum_{i=3}^{v_H} \binom{n}{i, v_H - i, v_H - i} e_d(i) \\
&= \sum_{i=3}^{v_H} \binom{n}{2v_H - i}\binom{2v_H - 1}{i}\binom{2v_H - 2i}{v_H - 1} e_d(i) \\
&= \mathcal{O}\left(\sum_{H'\subset H} n^{2v_H - v_{H'}}\right)
\end{aligned}$$

So, one deduces that, as $n \to \infty$:

$$\begin{aligned}\mathrm{Var}[N_H] &\asymp \sum_{\substack{H' \subset H \\ e_{H'} > 0}} n^{2v_H - v_{H'}} p_n^{2e_H - e_{H'}} (1 - p_n^{e_{H'}}) \\ &\lessapprox (1 - p_n^{e_H}) \sum_{\substack{H' \subset H \\ e_{H'} > 1}} n^{2v_H - v_{H'}} p_n^{2e_H - e_{H'}} (1 - p_n^{e_{H'}}) \\ &\asymp (1 - p_n^{e_H}) \mathbb{E}[N_H] + o(1).\end{aligned}$$

Conclusion

Remark 3.2.4. It is mentioned in Ruciński (1988, theorem 1) that the Poisson approximation is easier to derive than normal approximation as for a long time, asymptotic normality of that type of U-statistics has been proven using method of moments.

In Coulson et al. (2016), it is more or less the same theorem as previously but applied astutely.

More recently, using the Cramér-Wold device, Coulson et al. (2016) derived Poisson approximation of subgraph counts for model with conditional independence.

3.3 Normal approximation

The asymptotic normality of subgraph count in Erdös-Rényi model is well-known, as well as the convergence rate Janson and Nowicki (1991). The asymptotic normality of subgraph counting of random graphs is, in fact, a mere application of normal approximation of U-Statistics and homogeneous sums (Janson and Nowicki, 1991; De Jong, 1990). It is one of the main applications of discrete Stein-Malliavin method, but surprisingly the proofs for that problem of estimation has not been tackled until recently (Krokowski et al., 2017; Privault and Serafin, 2018; Privault et al., 2020; Privault and Serafin, 2022). In the light of those papers, it seems one can find such results of convergence rates for various statistics on random graphs. More involved applications are extensively described in Janson and Nowicki (1991). Interestingly enough, recently, Döbler, Kasprzak, and Peccati (2022) expresses the condition for asymptotic normality in terms of contractions which are of better use than moments in that particular application of normal approximations of U-Statistics. Surprisingly, the works for subgraph counting in random hypergraphs are even scarcer (De Jong, 1996), whereas the results are slightly different in that general case. There are many extensions that revolve around the definition of a random graph as a sequence of independent random variables, for example a clique complex of Bernoulli random graphs. That section also aims at putting forward new examples of applications of approximation theorems for U-Statistics.
Historically, normal approximation for subgraph counting had been dealt with method of moments (Ruciński, 1988), which requires tedious computations, but is quite adapted to this application. Later, Barbour et al. (1989) used Stein's method to derive convergence rates of the number of subgraph counting in random graphs, only in the Wasserstein-1 distance as pointed out by Röllin (2022). The latter derived Kolmogorov bounds for the normal approximation of the number of triangles in the Erdös-Rényi model. It is also an application of Krokowski et al. (2017). We present here a different approach using our partial fourth moment theorem for normal approximation in the Wasserstein-1 distance. We note that Privault and Serafin

(2018); Privault et al. (2020) manages to obtain Kolmogorov bounds in full generality leveraging integration by parts given by Malliavin calculus.

The overarching goal is to determine threshold functions for normal approximation of subghyperraph counts. We follow DeJong's approach to the problem, that leverages the Hoeffding decomposition. That would prove useful for more general hypergraph statistics.
We take the special case of r-uniform random hypergraphs for a given $r \geq 2$.

Example 3.3.1 (Subgraph counting in Erdös-Rényi model (De Jong, 1996))**.** The number of subhypergraphs of $\mathbb{G}^{(3)}(n, p_n)$ isomorphic to H is

$$M_H = \sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \prod_{\alpha \in I} \hat{X}_\alpha, \tag{3.3.1}$$

where $\simeq$ stands for hypergraph isomorphism. We denote by $\mathrm{Aut}(H)$ the automorphism group of H, that is the set of permutation on vertices. For $\sigma \in \mathrm{Aut}(H)$, $(x, y, z) \in E(H)$ if and only if $(\sigma(x), \sigma(y), \sigma(z)) \in E(H)$. The random variable M_H has a finite Hoeffding decomposition (De Jong, 1996, p.11(115)). Since $X_\alpha = p_n + (\hat{X}_\alpha - p_n)$, M_H admits the decomposition:

$$M_H = \sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \sum_{J \subseteq I} p_n^{|I|-|J|} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n), \tag{3.3.2}$$

where the summation extends over all subsets J of I. By interchanging the sums, we find the chaotic decomposition of $M_H - \mathbb{E}[M_H]$ that is:

$$\begin{aligned}
M_H - \mathbb{E}[M_H] &= \sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \sum_{\substack{J \subseteq I \\ J \neq \emptyset}} p_n^{|I|-|J|} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n), \\
&= \sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \sum_{j=1}^{|E(H)|} p_n^{|E(H)|-j} \sum_{\substack{J \subset I \\ |J| = j}} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n) \\
&= \sum_{j=1}^{|E(H)|} p_n^{|E(H)|-j} \sum_{|J|=j} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n) \left(\sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H, I \supseteq J}} 1 \right) \\
&= \sum_{j=1}^{|E(H)|} \pi_j(N_H),
\end{aligned}$$

where:

$$\pi_k(N_H) = p_n^{|E(H)|-j} \sum_{|J|=j} \left(\sum_{\substack{I \in \binom{[n]}{3s} \\ I \simeq H, I \supseteq J}} 1 \right) \prod_{\alpha \in J} \hat{Y}_\alpha \tag{3.3.3}$$

with $\hat{Y}_\alpha$ is the centered version of $\hat{X}_\alpha$ for all α hyperedges of K_n. We note that the decomposition above corresponds to the Hoeffding decomposition of the U-statistics with

$$W_J \propto \left(\sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H, I \supseteq J}} 1 \right) \prod_{\alpha \in J} \hat{Y}_\alpha. \tag{3.3.4}$$

We proceed in the same manner in $\mathbb{T}^{(3)}(n, q_n, p_n)$. Let N_H the number of pattern isomorphic to H

$$N_H = \sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \prod_{\alpha \in I} X_\alpha, \tag{3.3.5}$$

Here, $(X_\alpha)_{\alpha \in \binom{[n]}{3}}$ is a sequence of conditionally independent Bernoulli random variables given $Z = \mathbb{G}(n, q_n)$. The chaos decomposition yields:

$$\begin{aligned} N_H &= \sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \sum_{J \subseteq I} \prod_{\beta \in I \setminus J} \mathbb{E}\left[X_\beta \,|\, \mathbb{G}(n, q_n)\right] \prod_{\alpha \in J} (X_\alpha - \mathbb{E}\left[X_\alpha \,|\, \mathbb{G}(n, q_n)\right]) \\ &= \sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \sum_{J \subseteq I} p_n^{|I|-|J|} \mathbb{1}_{\{(I \setminus J)^{(2)} \subset \mathbb{G}(n, q_n)\}} \prod_{\alpha \in J} (X_\alpha - \mathbb{E}\left[X_\alpha \,|\, \mathbb{G}(n, q_n)\right]). \end{aligned} \tag{3.3.6}$$

Hence, $N_H - \mathbb{E}\left[N_H \,|\, \mathbb{G}(n, q_n)\right]$ reads off:

$$\sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H}} \sum_{\emptyset \neq J \subseteq I} p_n^{|I|-|J|} \mathbb{1}_{\{(I \setminus J)^{(2)} \subset \mathbb{G}(n, q_n)\}} \prod_{\alpha \in J} (X_\alpha - \mathbb{E}\left[X_\alpha \,|\, \mathbb{G}(n, q_n)\right]). \tag{3.3.7}$$

The corresponding degenerate U-statistics in the decomposition are given for $J \subset \binom{[n]}{3}$ by

$$W_J \propto \left(\sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H, I \supseteq J}} 1 \right) \prod_{\alpha \in J} Y_\alpha, \tag{3.3.8}$$

where Y_α is the centered version of X_α given $\mathbb{G}(n, q_n)$.

Various quantitative limit theorems exist in the literature. Ours leverage the fourth moment phenomenon.

The ancestor of fourth moment theorem was introduced by De Jong (1990), and was applied to problem of subgraph counting (De Jong, 1996). More recently, Bhattacharya et al. (2022) derived bounds in Wasserstein distance for the analogous problem of motif estimation. The fourth moment phenomenon that arise in there is the convergence of the fourth cumulant of the normalized (reduced) graph statistics to the fourth cumulant of the standard reduced normal distribution, which is 3. However, in the case of the subgraph counting, it seems that the appropriate bound is an ersatz of the fourth cumulant as it has been shown that the convergence of the fourth cumulant is not enough to prove asymptotic normality for functionals of Rademacher random variables (Döbler et al., 2019).

Repeat maximal influence

The index set of the random variables is here, the discrete set of (hyperedges).
We look for an upper bound of $|\mathbb{E}[W_I W_J W_K W_L]|$ for $I, J, K, L \in \binom{[n]}{r}$. That will give us the convergence rate.

Theorem 3.3.2 — Central Limit theorem De Jong (1996). *Let $X = (X_1, X_2, \dots)$ a sequence of independent random variables. Suppose Z is $\mathcal{F}_{[n]}$ measurable, with finite Hoeffding decomposition such that $\operatorname{Var}[Z] = 1$ (Z is a degenerate U-statistics of order d):*

$$Z = \sum_{I \subset [n], |I| \leq d} W_I, \tag{3.3.9}$$

satisfies the two conditions:

1. $\max_{i \in [\![1,n]\!]} \sum_{I \ni i} \mathbb{E}[W_I^2] \to 0$, *for* $n \to +\infty$;
2. $\sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]| \to 0$ *for* $n \to +\infty$.

De Jong applied that theorem to subgraph counting in random hypergraphs, with edges defined as random subsets of a vertex set.

Remark 3.3.3. If p_n is bounded away from 0, the condition is not sharp.

The first condition is not very severe. It is Lindeberg-Feller type condition as for $d = 1$ it is the crucial assumption in the Lindeberg-Feller central limit theorem (see for example theorem 5.12 in Kallenberg (2002)). In the homogeneous case where $|I| = d$, the fourth moment condition is sufficient, but otherwise it is not.

3.4 Application of Malliavin-Stein's method

We generalize the main theorem of Bhattacharya et al. (2022) for random multilinear forms in Bernoulli variables by using Proposition 2.7.10.

Theorem 3.4.1 *Let H a hypergraph without isolated vertices. Then, we have*

$$d_W(\bar{M}_H, \mathcal{N}(0,1)) \lesssim \left((1-p_n)^{e_H} \min_{\substack{H' \subset H \\ e_{H'} > 1}} \{ n^{v_H} p_n^{e_H} \} \right)^{-1/2} \tag{3.4.1}$$

and

$$d_W(\tilde{N}_H, \mathcal{N}(0,1)) \lesssim \left((1-p_n)^{e_H} \min_{\substack{H' \subset H \\ e_{H'} > 1}} \{ n^{v_H} p_n^{e_H} q_n^{e_H^{(2)}} \} \right)^{-1/2}, \tag{3.4.2}$$

where $e_H^{(2)}$ is the number of edges included in the hyperedges of H.

Proof of theorem 3.4.1. We proceed as in the previous proof given the decomposition of M_H

and borrow the previous notations. The (EGF) assumption holds. By conditional independence of $(X_\alpha)_\alpha$, we have:

$$\frac{\mathbb{E}[W_I^2]\mathbb{E}[W_J^2]}{\mathbb{E}[W_I^2W_J^2]} \propto \frac{\mathbb{E}\left[\prod_{\alpha\in I}\mathbb{E}[Y_\alpha^2|Z]\prod_{\alpha\in J}\mathbb{E}[Y_\alpha^2|Z]\right]}{\prod_{I\setminus J}\mathbb{E}[Y_\alpha^2]\prod_{J\setminus I}\mathbb{E}[Y_\alpha^2]\prod_{I\cap J}\mathbb{E}[Y_\alpha^2]}$$
$$= (p_n(1-p_n))^{|I|+|J|-|I\cup J|}\, q_n^{|I^{(2)}|+|J^{(2)}|-|I^{(2)}\cup J^{(2)}|} \leq 1.$$

Let us note that for all a, $W_{I\setminus\{a\}}$ is non-zero with the definition of $M_H - \mathbb{E}\left[M_H|Z\right]$. Let $W_I = w_I\prod_{i\in I}X_i$, then for $a\in I\cap J$:

$$\mathbb{E}\left[W_IW_J|\mathcal{G}_a\right] = w_Iw_J\prod_{i\in I\setminus\{a\}}Y_i\prod_{j\in J\setminus\{a\}}Y_j\mathbb{E}\left[Y_a^2|Z\right]$$
$$= \frac{w_Iw_J}{w_{I\setminus\{a\}}w_{J\setminus\{a\}}}\mathbb{E}\left[Y_a^2|Z\right]W_{I\setminus\{a\}}W_{J\setminus\{a\}}$$
$$= C_{I,J,a}W_{I\setminus\{a\}}W_{J\setminus\{a\}},$$

with

$$C_{I,J,a} = \frac{w_Iw_J}{w_{I\setminus\{a\}}w_{J\setminus\{a\}}}\mathbb{E}\left[Y_a^2|Z\right] < +\infty\ \mathbb{P}\text{-almost surely.}$$

We are then left to upper bound the quantity:

$$\sum_{(I,J,K,L)\text{ connected}}|\mathbb{E}[W_IW_JW_KW_L]|$$
$$= \sum_{(I,J,K,L)\text{ connected}}\left|\mathbb{E}\left[\mathbb{E}\left[\prod_{\alpha\in I}Y_\alpha\prod_{\alpha\in J}Y_\alpha\prod_{\alpha\in K}Y_\alpha\prod_{\alpha\in L}Y_\alpha|Z\right]\right]\right|.$$

The terms are non-zero if only if α lies in at least two elements of the quadruple, i.e. if α does not lie in $I\setminus(J\cup K\cup L)$, etc. Then, the number of non-zero terms is $I\cup J\cup K\cup L$. We recall that:

$$\mathbb{E}[Y_\alpha|Z] = 0$$
$$\mathbb{E}[Y_\alpha^2|Z] = p_n(1-p_n)\mathbb{1}_{\{\alpha^{(1)}\in Z\}}\prod_{i=1}^{3}\mathbb{1}_{\{\alpha^{(i)}\in Z\}}$$
$$\mathbb{E}[Y_\alpha^3|Z] = p_n(1-p_n)(1-2p_n)\prod_{i=1}^{3}\mathbb{1}_{\{\alpha^{(i)}\in Z\}} \lesssim p_n(1-p_n)^2\prod_{i=1}^{3}\mathbb{1}_{\{\alpha^{(i)}\in Z\}}$$
$$\mathbb{E}[Y_\alpha^4|Z] = p_n(1-p_n)(1-3p_n(1-p_n))\prod_{i=1}^{3}\mathbb{1}_{\{\alpha^{(i)}\in Z\}} \lesssim p_n(1-p_n)\prod_{i=1}^{3}\mathbb{1}_{\{\alpha^{(i)}\in Z\}}.$$

Thus,

$$\sum_{(I,J,K,L)\text{ connected}}|\mathbb{E}[W_IW_JW_KW_L]|$$
$$\lesssim \sum_{(I,J,K,L)\text{ connected}}(p_n(1-p_n))^{|I\cup J\cup K\cup L|}\, q_n^{|I^{(2)}\cup J^{(2)}\cup K^{(2)}\cup L^{(2)}|}$$

Now, we remark I, J, K, L are respectively isomorphic to A, B, C, D subhypergraphs of H. We say that $I \sim_H A$, etc. Hence, we can sum first over (A, B, C, D). and then over all the quadruples (I, J, K, L) whose components are respectively isomorphic to the ones of the fixed quadruple (A, B, C, D). We shall write:

$$\sum_{I,J,K,L} \cdot = \sum_{A,B,C,D} \sum_{\substack{I \sim_H A, J, \sim_H B \\ K \sim_H C, L \sim_H D}} \cdot := \sum_{A,B,C,D} \sum_{I,J,K,L}^{*A,B,C,D} \cdot$$

$v(A)$ denotes the number of vertices in A. We have that $|\{I, J, K, L \in \binom{[n]}{r} : I \simeq A, J \simeq B, K \simeq C, L \simeq D\}|$ is bounded by the number of collection of vertices of cardinal $v(A \cup B \cup C \cup D)$. By a counting argument, we see that is of order $n^{v(A \cup B \cup C \cup D)}$. Because (I, J, K, L) is connected and copies of subhypergraphs of H, we also have that $|I \cup J \cup K \cup L| = |A \cup B \cup C \cup D|$ and $|I^{(2)} \cup J^{(2)} \cup K^{(2)} \cup L^{(2)}| = |A^{(2)} \cup B^{(2)} \cup C^{(2)} \cup D^{(2)}|$. Hence,

$$\begin{aligned}
\sum_{(I,J,K,L) \text{ connected}} & |\mathbb{E}[W_I W_J W_K W_L]| \\
&= \sum_{A,B,C,D} \sum_{(I,J,K,L) \text{ connected}}^{*A,B,C,D} p_n^{|A \cup B \cup C \cup D|} q_n^{|A^{(2)} \cup B^{(2)} \cup C^{(2)} \cup D^{(2)}|} \\
&\leq \sum_{A,B,C,D} n^{v(A \cup B \cup C \cup D)} p_n^{|A \cup B \cup C \cup D|} q_n^{|A^{(2)} \cup B^{(2)} \cup C^{(2)} \cup D^{(2)}|}
\end{aligned}$$

Let us bound the variance of $N_H - \mathbb{E}[N_H | \mathbb{G}(n, q_n)]$:

$$\begin{aligned}
\mathrm{Var}^2[N_H - \mathbb{E}[N_H | \mathbb{G}(n, q_n)]] &= \left(\sum_{I \cap J \neq \emptyset} \mathbb{E}[W_I W_J] \right)^2 = \sum_{I \cap J \neq \emptyset} \left(\mathbb{E}[W_I^2] + \mathbb{E}[W_J^2] \right)^2 \\
&= \frac{1}{2^2} \sum_{A,B \subset H} \left(\sum_I^{*A} \mathbb{E}[W_I^2] + \sum_J^{*B} \mathbb{E}[W_J^2] \right)^2.
\end{aligned}$$

For a fixed connected quadruple (A, B, C, D),

$$\begin{aligned}
\mathrm{Var}^2&[N_H - \mathbb{E}[N_H | \mathbb{G}(n, q_n)]] \\
&\geq \frac{1}{16} \left(\sum_I^{*A} \mathbb{E}[W_I^2] + \sum_J^{*B} \mathbb{E}[W_J^2] + \sum_K^{*C} \mathbb{E}[W_K^2] + \sum_L^{*D} \mathbb{E}[W_L^2] \right)^2 \\
&\geq \frac{1}{16} \left(\sum_I^{*A} \mathbb{E}[W_I^2] \times \sum_J^{*B} \mathbb{E}[W_J^2] \times \sum_K^{*C} \mathbb{E}[W_K^2] \times \sum_L^{*D} \mathbb{E}[W_L^2] \right)^{1/2},
\end{aligned}$$

applying repeatedly the inequality $a^2+b^2 \geq 2ab$. Then we use that $\mathbb{E}[W_I^2] = q_n^{|I^{(2)}|}(1-p_n)^{|I|} p_n^{|I|} = q_n^{|A^{(2)}|}(1-p_n)^{|A|} p_n^{|A|}$, so

$$\sum_I^{*A} \mathbb{E}[W_I^2] = \sum_I^{*A} (1-p_n)^{|A|} p_n^{|A|} q_n^{|A^{(2)}|} = n^{v(A)} (1-p_n)^{|A|} p_n^{|A|} q_n^{|A^{(2)}|}$$

In particular, one has:

$$\mathrm{Var}^2[N_H - \mathbb{E}\left[N_H \middle| \mathbb{G}(n,q_n)\right]] \geq \frac{1}{16}\left(n^{v^*(A,B,C,D)}(p_n(1-p_n))^{e^*(A,B,C,D)}q_n^{e^{(2)*}(A,B,C,D)}\right)^{1/2} \quad (3.4.3)$$

where $v^*(A,B,C,D) = v(A)+v(B)+v(C)+v(D)$ and $e^*(A,B,C,D) = |A|+|B|+|C|+|D|$ and $e^{(2)*}(A,B,C,D) = |A^{(2)}|+|B^{(2)}|+|C^{(2)}|+|D^{(2)}|$. It yields the result. Using the lemma 9 of De Jong (1996), (3.4.2) follows. The first result for M_H is obtained with $q_n = 1$. □

We now explore general U-Statistics on random hypergraphs and their applications in the scope of the taxonomy of random hypergraphs presented in the first chapters. There is already study of asymptotic normality in the subgraph count in Stochastic Block Model in Janson and Nowicki (1991) which is a random graph with conditionally independent edges, without the associated threshold functions. We propose to extend that, and consider random hypergraphs generated by hypergraphons.

Example 3.4.2. Subgraph counting in Erdös-Rényi egdes + Rips
Conditionally on drawing the edges, select every triangle with probability p_n. We want to count the number of hypergraph motifs (or triangles for example). The random variables are independent conditionally given the graph $\mathbb{G}(n,q_n)$ following $\mathrm{Ber}(q_n)$. The estimator reads off:

$$N(H,\mathbb{G}(n,q_n,p_n)) := \frac{1}{|Aut(H)|}\sum_{\boldsymbol{e}\in\binom{[n]}{2}} \mathbb{1}_{\{e_1,\}} \prod_{\alpha\in E(H)} \hat{X}_{e_\alpha}, \quad (3.4.4)$$

where $\hat{X}_{e_\alpha} \sim \mathcal{B}(p_n)$ conditionally i.i.d. given $(X_{e_{u,v}}, X_{e_{v,w}}, X_{e_{w,u}})$ for $\alpha = (u,v,w)$, having $X_\beta \sim \mathcal{B}(q_n)$, with e a function that sends the hyperedges of H into $([n], \binom{[n]}{2})$. We can also write like Privault and Serafin (2018):

$$N_H \propto_H \sum_{b_1,\dots,b_{e_H}} \mathbb{1}_{\{(b_1,\dots,b_{e_H})\in E(H)\}}(Y_{b_1}+1)\dots(Y_{b_{e_H}}+1) \quad (3.4.5)$$

where Y_e is a Rademacher random variable with parameter p_n. $(Y_e)_{e\in\mathbb{N}}$ is a conditionally independent Rademacher sequence given $\mathbb{G}(n,q_n)$. We first compute the expectation and variance of N_H. If $p_n = 1$, we know from Elek and Szegedy (2012, p.12 example 2.) that the almost sure limit of the homomorphism density of H in the random hypergraph is $q_n^{|E^{(2)}(H)|}$ where $E^{(2)}(H)$ is the set of edges in the simplicial complex associated to H. As a consequence, because of concentration results,

$$\mathbb{E}[N_H] = n^{|V(H)|}q_n^{|E^{(2)}(H)|}.$$

More generally, we have:

$$\mathbb{E}[N_H] \propto p_n^{|E(H)|}q_n^{|E^{(2)}(H)|}.$$

Let us compute the variance of the estimator. To see this, recall that if: $|e^{\mathbf{1}} \cap e^{\mathbf{2}}| = K$, and $|(e^{\mathbf{1}} \cap e^{\mathbf{2}})^{(2)}| = J$ (the number of edges in the intersection of hyperedges) then:

$$\begin{aligned}
\mathrm{Cov}\Big(\prod_{\alpha\in E(H)} X_{e^{\mathbf{1}}_\alpha}, \prod_{\alpha\in E(H)} X_{e^{\mathbf{2}}_\alpha}\Big) &= \mathbb{E}\left[\prod_{\alpha\in E(H)} X_{e^{\mathbf{1}}_\alpha} X_{e^{\mathbf{2}}_\alpha}\right] - \mathbb{E}\Big[\prod_{\alpha\in E(H)} X_{e^{\mathbf{1}}_\alpha}\Big]\mathbb{E}\Big[\prod_{\alpha\in E(H)} X_{e^{\mathbf{2}}_\alpha}\Big] \\
&= \mathbb{E}\left[\prod_{\alpha\in E(H)} X_{e^{\mathbf{1}}_\alpha} X_{e^{\mathbf{2}}_\alpha}\right] - p_n^{2|E(H)|} q_n^{2|E^{(2)}(H)|} \\
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{\alpha\in E(H)} X_{e^{\mathbf{1}}_\alpha} X_{e^{\mathbf{2}}_\alpha} \mid \mathbb{G}(n, q_n)\right]\right] - p_n^{2|E(H)|} q_n^{2|E^{(2)}(H)|} \\
&= q_n^{2|E^{(2)}(H)|-J} p_n^{2|E(H)|-K} - q_n^{2|E^{(2)}(H)|} p_n^{2|E(H)|}
\end{aligned}$$

So, using that for each $H' \subset H$, there are $\Theta(n^{v'_H} n^{2(v_H - v_{H'})}) = \Theta(n^{2v_H - v_{H'}})$ pairs $(H^{\mathbf{1}}, H^{\mathbf{2}})$ copies of H in the complete graph K_n with $H^{\mathbf{1}} \cap H^{\mathbf{2}}$ isomorphic to H' (here the hyperedge sets).

$$\begin{aligned}
\mathrm{Var}[N_H] &\propto_H \sum_{e^{\mathbf{1}}, e^{\mathbf{2}} \in \binom{[n]}{2}} \mathrm{Cov}\left(\prod_{\alpha\in E(H)} X_{e^{\mathbf{1}}_\alpha}, \prod_{\alpha\in E(H)} X_{e^{\mathbf{2}}_\alpha}\right) \\
&= \sum_{\substack{e^{\mathbf{1}}, e^{\mathbf{2}} \in \binom{[n]}{2} \\ e^{\mathbf{1}} \cap e^{\mathbf{2}} \neq \emptyset}} \mathrm{Cov}\left(\prod_{\alpha\in E(H)} X_{e^{\mathbf{1}}_\alpha}, \prod_{\alpha\in E(H)} X_{e^{\mathbf{2}}_\alpha}\right) \\
&= \sum_{K=1}^{|E(H)|} \sum_{J=1}^{|E^{(2)}(H)|} \sum_{\substack{e^{\mathbf{1}}, e^{\mathbf{2}} \in \binom{[n]}{2} \\ |e^{\mathbf{1}} \cap e^{\mathbf{2}}| = K \\ |(e^{\mathbf{1}} \cap e^{\mathbf{2}})^{(2)}| = J}} \left(q_n^{2|E^{(2)}(H)|-J} p_n^{2|E(H)|-K} - q_n^{2|E^{(2)}(H)|} p_n^{2|E(H)|}\right) \\
&\asymp \sum_{\substack{H' \subset H \\ e_{H'} \geq 1}} n^{2v_H - v_{H'}} q_n^{2|E^{(2)}(H)| - |E^{(2)}(H')|} p_n^{2|E(H)| - |E(H')|} - q_n^{2|E^{(2)}(H)|} p_n^{2|E(H)|} \\
&= \sum_{\substack{H' \subset H \\ e_{H'} \geq 1}} n^{2v_H - v_{H'}} q_n^{2|E^{(2)}(H)| - |E^{(2)}(H')|} p_n^{2|E(H)| - |E(H')|} \left(1 - q_n^{|E^{(2)}(H')|} p_n^{|E(H')|}\right)
\end{aligned}$$

We consider the case where $p_n \to 0$ and/or $q_n \to 0$ for $n \to +\infty$. So asymptotically (extension of Lemma 3.5. in Janson et al. (2000) to random hypergraphs), we have:

$$\mathrm{Var}[N_H] \asymp_H (1 - p_n q_n) \max_{\substack{H' \subset H \\ e_{H'} \geq 1}} n^{2v_H - v_{H'}} q_n^{2|E^{(2)}(H)| - |E^{(2)}(H')|} p_n^{2|E(H)| - |E(H')|}. \tag{3.4.6}$$

In the view of this expression, p_n and q_n share the same role for the convergence.

We need a modified Hoeffding decomposition for such statistics.

3.5 A modified Hoeffding decomposition

We shall write the sequence of conditionally independent random variables $\mathsf{X} = (\hat{X}_\alpha, \ldots, X_\beta, \ldots)_{\alpha,\beta \in A}$. Since $\sigma(Z) = \sigma(\hat{X}_a, a \in A)$, we adopt a sightly different approach by leveraging the Malliavin-Dirichlet structure in Decreusefond and Halconruy (2019), whose underlying Markov process is the classical Glauber dynamics starting from X. The structure is similar to ours if the expectation are taken conditionally with respect to Z for the operators and conditional distributions in section 2.2. The following lemma shows the commutation relation on Dom D in that Malliavin framework.

Lemma 3.5.1 *For $F \in L^2(E_A)$ and $\alpha, \beta \in A$,*

1.

$$\mathbb{E}\left[\mathbb{E}\left[F\,|\,\hat{X}^{\{\beta\}}, X\right]|\,X^{\{\alpha\}}, \hat{X}\right] = \mathbb{E}\left[\mathbb{E}\left[F\,|\,X^{\{\alpha\}}, \hat{X}\right]|\,X, \hat{X}^{\{\beta\}}\right]; \tag{3.5.1}$$

2. for $\alpha \neq \beta$

$$\mathbb{E}\left[\mathbb{E}\left[F\,|\,\hat{X}^{\{\alpha\}}\right]|\,\hat{X}^{\{\beta\}}\right] = \mathbb{E}\left[\mathbb{E}\left[F\,|\,\hat{X}^{\{\beta\}}\right]|\,\hat{X}^{\{\alpha\}}\right]. \tag{3.5.2}$$

Proof. Since $X_\beta = \frac{\mathbb{1}_{\{U_\beta \le p_n\}}}{p_n}\hat{X}_\beta$, we have:

$$\mathbb{E}\left[\hat{X}_\beta\,|\,\hat{X}^{\{\beta\}}, X\right] = \mathbb{E}\left[X_\beta\,|\,\hat{X}^{\{\beta\}}, X\right] = X_\beta,$$

where U_β is a uniform random variable independent of the rest for $\beta \in A$. Thus, we get that $\mathbb{E}\left[F\,|\,\hat{X}^{\{\beta\}}, X\right] = F$, hence the relations using lemma 3.5.1. □

Those commutation relations of lemma 3.5.1 and lemma 2.2.4 entail a modified Hoeffding decomposition of N_H similar to (3.3.6) that this time the conditional mean has the chaotic decomposition:

$$\mathbb{E}\left[M_H\,|\,Z\right] = \sum_{j=1}^{|E(H)|} (p_n q_n^3)^{|E(H)|-j} \sum_{|J|=j} \prod_{\alpha \in J} \hat{Y}_\alpha \left(\sum_{\substack{I \in \binom{[n]}{r} \\ I \simeq H, I \supseteq J}} 1 \right)$$

where $\hat{Y}_\alpha = \hat{X}_\alpha - \mathbb{E}[\hat{X}_\alpha]$ with $\hat{X}_\alpha = \mathbb{E}\left[X_\alpha\,|\,Z\right]$. Thus,

$$\begin{aligned}\pi_j(N_H) = p_n^{|E(H)|-j} \sum_{|J|=j} \left(\sum_{\substack{I \in \binom{[n]}{3} \\ I \simeq H, I \supseteq J}} 1 \right) \prod_{\alpha \in J} \tilde{Y}_\alpha \\ + (p_n q_n^3)^{|E(H)|-j} \sum_{|J|=j} \prod_{\alpha \in J} \hat{Y}_\alpha \left(\sum_{\substack{I \in \binom{[n]}{r} \\ I \simeq H, I \supseteq J}} 1 \right)\end{aligned} \tag{3.5.3}$$

All the previous bounds are still valid in that context by taking out Z of the conditioning and can be applied given that chaotic decomposition.

Theorem 3.5.2 *Let H a hypergraph without isolated vertices. Then, let $p_n \xrightarrow{n\to+\infty} 0$ and $q_n \xrightarrow{n\to+\infty} 0$:*

$$d_W(\bar{N}_H, \mathcal{N}(0,1)) \lesssim \left((1-p_n)^{e_H} \min_{\substack{H'\subset H \\ e_{H'}>1}} \{ n^{v_H} p_n^{e_H} q_n^{e_H^{(2)}} \} \right)^{-1/2}. \quad (3.5.4)$$

Proof of theorem 3.5.2. We follow the same lines as the previous proof, with the difference that $\pi_0(N_H) = \mathbb{E}[F]$, leading to the quantitative central limit theorem. □

While in Kaur and Röllin (2021); Temčinas et al. (2022), the probability of keeping a hyperedge does not depend on the number of vertices, we let p_n tend to 0. As a consequence, we can state thresholds for subhypergraph containment that complement the ones in (Janson et al., 2000, p.61). As done in Zhang (2022); Kaur and Röllin (2021) for random graphs, it should be possible to derive with our method the convergence rates considering an arbitrary exchangeable random hypergraph generated by a hypergraphon, the analog of graphon in graph limit theory.

3.6 Limit theorems for homomorphism densities of random hypergraphs

See Delmas et al. (2021).
A result for bivariate homomorphism densities (for simple patterns) by Fang and Röllin (2015).
We want to derive bounds in probability distances for smooth functions (ideally Kolmogorov bounds) as to get a test for confidence interval.
We still use the derived limit theorem found earlier.

Part II

Invertibility of functionals of marked point processes

The starting point for our second work is the existence of a new construction of Hawkes processes as a functional of an underlying Poisson process.
Based on the extension of the invertibility framework introduced by Üstünel (2009) and recently in the Poisson space by Coutin and Decreusefond (2023), we find an entropic criterion under which a random change of mark of marked Poisson point process is invertible using the Girsanov theorem.

Chapter 4

Invertibility of functionals of Poisson measures

The Skorokhod theorem on invariance of measures Skorokhod (1957) gives the density with respect to Poisson measures of deterministic shifts of configuration. This theorem has an extension to random transformation of marks of a marked Poisson process N. In that case, the key property is quasi-invariance with respect to anticipative transformation Albeverio and Smorodina (2006); Privault (1996) which is given by a classical application of the Girsanov theorem. The specificity of the Poisson setting is the identification of $\mathfrak{M}$ the space of random measures and the configuration space (see the detailed definition below). Namely, let $\mathbf{\Gamma}$ the transformation applied on a Poisson measure N under a probability measure π on $\mathfrak{M}$. As N can also be considered a marked point process, $\mathbf{\Gamma}$ transposes to a random transformation of an element $\omega \in \mathfrak{N}$, namely the marks Z_n of $\omega = (T_n, Z_n)_{n\in\mathbb{N}}$ are changed to $\gamma(T_n, Z_n)$ for $n \in \mathbb{N}$. Under some assumptions on $\mathbf{\Gamma}$ detailed in the following, there exists a probability measure π' under which both marked point processes $\mathbf{\Gamma}(N)$ and N share the same law. The Girsanov theorem also gives the proof of the existence of weak solution of stochastic differential equation driven by Brownian motion or Poisson measures or both under weak assumptions. The stochastic differential equation (SDE for short) that we consider in this chapter, is of the form:

$$\begin{cases} Y(0) &= 0 \\ \mathrm{d}Y(t) &= \int b(Y(t_-), x)\ \mathrm{d}N(t, x). \end{cases} \tag{4.0.1}$$

where N is a Poisson measure with control measure $\nu(\mathrm{d}x) \otimes \mathrm{d}t$. Denote by $\mathfrak{M}$ the space of random counting processes on $\mathbb{R}^+ \times \mathbb{R}$. By identification of random measures and random counting processes, N is an element of $\mathfrak{M}$. Let π a probability on the space $\mathfrak{M}$ such that N is the canonical counting process on it, namely $\mathrm{Id}_{\mathfrak{M}}$. Let $\mathbb{D}$ the space of cadlag functionals. Consider the map

$$\begin{aligned} V\ :\ \mathfrak{M} &\longrightarrow \mathbb{D} \\ N &\longmapsto \left(t \mapsto \int_0^t \int x\ \mathrm{d}N(s, x) \right). \end{aligned} \tag{4.0.2}$$

Let $\mathbf{\Upsilon}$ a perturbation of the sample path and by identification of $\omega \in \mathfrak{N}$ such that:

$$\mathbf{\Upsilon}(\omega)(t, z) = (T_n, b(\sum_{n\in\mathbb{N}:T_n<t} Z_n(\omega), z)),$$

Not an esay introduction without introducing properly marked processes

A solution of the SDE (4.0.1) is

$$Y = V \circ \boldsymbol{\Upsilon}(N).$$

It induces a map $\mathbf{Y}$ on $\mathbb{D}$ such that $\mathbf{Y} \circ V(N) = Y$. Let $\mathbf{Z}$ a map on $\mathbb{D}$ of the form $V \circ \boldsymbol{\Gamma}$ for $\boldsymbol{\Gamma}$ some random transformation, we have:

$$\mathbf{Y} \circ \mathbf{Z} = \mathrm{Id}_{\mathbb{D}} \iff \boldsymbol{\Upsilon} \circ \boldsymbol{\Gamma}(N) = \mathrm{Id}_{\mathfrak{M}}.$$

Solving (4.0.1) revolves to invert the map $\boldsymbol{\Upsilon}$. That formulation of the problem puts the focus on the map $\boldsymbol{\Upsilon}$ on $\mathfrak{M}$ instead of $Y \in \mathbb{D}$. To the best of our knowledge, this approach has been seldom addressed in the literature.
The framework for invertibility of Poisson functionals was introduced in Coutin and Decreusefond (2023). The key is the quasi-invariance property from the Girsanov theorem. On the real line, the consequences of the quasi-invariance property have been studied with respect to "anticipative" (predictable) transformations in Privault (1996) and in the general case of metric spaces in Albeverio and Smorodina (2006). In the Wiener case, random non-adapted transformations of Brownian motion have been considered by several authors in the context of Malliavin calculus, cf. Üstünel and Zakai (2000) and references therein. The study of invertibility concerns perturbations of the sample path.

4.1 Preliminaries

In the following, let $\mathbb{M}$ a metric space and $\mathcal{B}(\mathbb{M})$ its Borel σ-algebra. We recall the construction of marked point processes which is a sequence of points drawn at random in $\mathbb{R}^+ \times \mathbb{M}$ in the configuration space $\mathfrak{N}$ endowed with the vague convergence, its Borel σ-algebra $\mathcal{B}(\mathfrak{N})$ and a probability measure μ.

Definition 4.1.1 (Marked point process)**.** A marked point process (MPP for short) is a sequence

$$\omega = (T_n(\omega), Z_n(\omega))_{n\in\mathbb{N}} \text{ of } \mathbb{R}^+ \times \mathbb{M} \text{ defined on } (\mathfrak{N}, \mathcal{B}(\mathfrak{N}), \mu)$$

such that for each $n \geq 1$, $T_n(\omega) < T_{n+1}(\omega)$, $T_n(\omega)$ tends to infinity μ-a.s. as n tends to infinity. The random variable T_n represents the n-th jump time and Z_n is the location associated to the n-th jump. It can be completely identified with a random counting measure ξ, viz.

$$\xi = \sum_{n\in\mathbb{N}:T_n(\omega)<\infty} \epsilon_{(T_n,Z_n)} \tag{4.1.1}$$

where $\epsilon_{(T_n,Z_n)}(\omega) = \epsilon_{(T_n(\omega),Z_n(\omega))}$ is the Dirac mass on the product space $((0,+\infty) \times \mathbb{M})$ at the point $(T_n(\omega), Z_n(\omega))$. We denote $\mathfrak{M}$ the space of counting random measures that can be written in such form. We endow this Lusin space with the vague convergence as a family of positive integer-valued Radon measures $(\xi(\omega,\cdot) : \omega \in \mathfrak{M})$ on $(\mathbb{R}^+ \times \mathbb{M}, \mathcal{B}(\mathbb{R}^+ \times \mathbb{M}))$. The random measure can be identified to a random counting process X. For $A \in \mathcal{B}(\mathbb{M})$, we shall write the process

$$X(\omega, t, A) = \xi(\omega, (0, t] \times A)$$

that counts the number of events on $[0, t]$ matching a mark belonging to the set A. We define the ground process.

$$X_g(\omega, t) = X(\omega, t, \mathbb{M}).$$

Note that ω can be retrieved as:

$$T_n(\omega) = \inf\{t \geq 0 : X_g(\omega, t) = n\}$$

$$(Z_n(\omega) \in A) = \bigcup_{K'=1}^{\infty} \bigcap_{K=K'}^{\infty} \bigcup_{k=1}^{\infty} \left(X_g(\omega, (k-1)/2^k = n-1, X_g(\omega, k/2^K, A) - X_g(\omega, (k-1)/2^K, A) = 1) \right).$$

That means that we can identify a sample-path X of $\mathfrak{M}$ with an element ω of $\mathfrak{N}$, and more generally elements of $\mathcal{B}(\mathfrak{M})$ and $\mathcal{B}(\mathfrak{N})$. We redefine μ as a probability on $(\mathfrak{M}, \mathcal{B}(\mathfrak{M}))$. Moreover, $\mathrm{d}X(s,z) = \mathrm{d}\xi(s,z)$.

Let $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ be an increasing and right-continuous family of sub-algebras of $\mathcal{B}(\mathfrak{M})$ such that each T_n is a stopping time, and each Z_n, is $\mathcal{F}_{T_n}$-measurable such that $\mathcal{B}(\mathfrak{M}) = \vee_{t \geq 0} \mathcal{F}_t$. We denote the following subsets of $\mathcal{P}(\mathcal{F})$ by:

- $\mathcal{P}^+(\mathcal{F})$ the set of predictable non-negative real-valued processes $(Y(t,z))_{(t,z) \in \mathbb{R}^+ \times \mathbb{M}}$ defined on $(\mathfrak{M}, \mathcal{F}_\infty, \mu)$;
- $\mathcal{P}^{++}(\mathcal{F})$ the set of predictable positive real-valued processes $(Y(t,z))_{(t,z) \in \mathbb{R}^+ \times \mathbb{M}}$ defined on $(\mathfrak{M}, \mathcal{F}_\infty, \mu)$.

We recall the version of Theorem 2.1 of Jacod (1975) for counting processes.

Theorem 4.1.2 *Let μ be a probability measure on $(\mathfrak{M}, \mathcal{F}_\infty)$. Let a random measure ξ on $\mathbb{R}^+ \times \mathbb{M}$ such that $t \mapsto \xi([0,t], \cdot)$ is adapted. Then there exists a unique predictable random measure denoted ξ^p such that*

$$\left(\xi([0,t], A) - \xi^p([0,t], A) \right)_{t \geq 0} \text{ is a } (\mathcal{F}, \mu)\text{-local martingale.}$$

The random measure ξ^p is called the dual predictable projection of ξ along $\mathcal{F}$. Equivalently, it means that for each $A \in \mathcal{B}(\mathbb{M})$, the process $t \mapsto \xi^p([0,t], A)$ is the so-called compensator of $t \mapsto \xi([0,t], A)$. We denote by $\tilde{\xi}$ the random measure $\xi - \xi^p$.

Definition 4.1.3. Let ξ a random measure. We introduce $\mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi)$ the space of all functions f in $\mathcal{P}(\mathcal{F})$ such that for each $t \in \mathbb{R}^+$,

$$\mathbf{E}_\mu \left[\int_0^t \int_{\mathbb{M}} |f(\cdot, s, z)| \xi(\mathrm{d}s, \mathrm{d}z) \right] < +\infty.$$

The Bochner integral $\int_0^t \int_{\mathbb{M}} f(\cdot, s, z) \xi(\mathrm{d}s, \mathrm{d}z)$ is a Lebesgue integral with respect to the measure $\xi(\omega,)$ for every $\omega \in \mathfrak{M}$ and is equal to the convergent sum

$$\int_0^t \int_{\mathbb{M}} f(\cdot, s, z) \xi(\mathrm{d}s, \mathrm{d}z) = \sum_{s \in (0,T] \cap \{s : \Delta X(s) \neq 0\}} f(\cdot, s, \Delta X(s)).$$

The stochastic integral of r with respect to ξ denoted by $\delta_\xi r$, is the process on $\mathbb{R}^+$

$$\delta_\xi r(t) = \int_0^t \int_{\mathbb{M}} r(s,z)\,\mathrm{d}\xi(s,z). \tag{4.1.2}$$

If $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi^p) \cap \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi)$, we define the integral with respect to the compensated random measure N by

$$\int_0^t \int f(s,z)\tilde{\xi}(\,\mathrm{d}s,\,\mathrm{d}z) = \int_0^t \int_{\mathbb{M}} f(s,z)\xi(\,\mathrm{d}s,\,\mathrm{d}z) - \int_0^t \int_{\mathbb{M}} f(s,z)\xi^p(\,\mathrm{d}s,\,\mathrm{d}z) \quad \mu\text{-a.s.} \tag{4.1.3}$$

Remark 4.1.4. In general, the stochastic integral cannot be divided as the difference of two integrals as in (4.1.3).

Proposition 4.1.5 *If $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi^p)$, then we have $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi)$ and for each $t \geq 0$,*

$$\mathbf{E}_\mu\left[\int_0^t \int_{\mathbb{M}} f(s,z)\xi(\,\mathrm{d}s,\,\mathrm{d}z)\right] = \mathbf{E}_\mu\left[\int_0^t \int_{\mathbb{M}} f(s,z)\xi^p(\,\mathrm{d}s,\,\mathrm{d}z)\right].$$

We define the Doléans-Dade exponential associated to a random measure ξ.

Definition 4.1.6. For $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \tilde{\xi})$, additionally supposed to be non-negative, the Doléans-Dade exponential Jacod (1975), denoted by $\mathcal{E}(\delta_{\tilde{\xi}} f)$, is defined as the solution of

$$M(t) = \int_0^t \int M(s^-) f(s,z)\,\mathrm{d}\tilde{\xi}(s,z),$$

explicitly given by

$$\tilde{\mathcal{E}}\left(\delta_{\tilde{\xi}} f\right)(t) = \exp\left(\int_0^t \int (1 - f(s,z))\,\mathrm{d}\tilde{\xi}(s,z)\right) \prod_{T_n \leq t} f((T_n, Z_n)) \tag{4.1.4}$$

In the case $r > -1$ and $r \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi^p)$, it is explicitly given by:

$$\begin{aligned} \mathcal{E}\left(\delta_{\tilde{\xi}} f\right)(t) &= \exp\left(\int_0^t \int \log(f(s,z))\,\mathrm{d}\xi(s,z) - \int_0^t \int (f(s,z) - 1)\,\mathrm{d}\xi^p(s,z)\right) \qquad (4.1.5) \\ &= \exp\left(\delta_{\tilde{\xi}}\Big(\log(f)\Big)(t) + \int_0^t \int \Big(\log\big(f(s,z)\big) - f(s,z) + 1\Big)\,\mathrm{d}\xi^p(s,z)\right). \end{aligned}$$

An important assumption in the remainder is that $\mathcal{F}$ is the minimal filtration to which X is adapted, viz. $\mathcal{F}$ is the filtration on $\mathfrak{M} \times \mathbb{R}^+$

$$\mathcal{F}^X = \sigma(X(t,A), t \in \mathbb{R}^+, A \in \mathcal{B}(\mathbb{M})). \tag{A1}$$

The following result is the converse of Theorem 4.1.2 (see Proposition 3.41 of Jacod (1979)).

Theorem 4.1.7 *For any nonnegative random measure ξ^p on $(0,\infty)\times\mathbb{M}$, there exists a unique probability measure on $(\mathfrak{M},\mathcal{F}^X)$ such that X has ξ^p as dual predictable projection.*

4.2 Background

4.3 Quasi-invariance

4.3.1 Random change of marks

Following the Bismut's approach Bismut (1983), we devise a perturbation γ of the jump sizes such that the map $\Gamma = \mathbf{\Gamma}(N)$ under π has the same law as N under π_ϕ, i.e. $\mathbf{\Gamma}_\#\pi = \pi_\phi$. We name the map triggers the perturbation a random change of marks in analogy to the random change of time for point process Coutin and Decreusefond (2023).

Definition 4.3.1 (Random transformation of mark)**.** On $\big(\mathfrak{M},\mathcal{F}_\infty,\mu\big)$, a random change of mark is a process $(\gamma(s,z),\,(s,z)\in\mathbb{R}^+\times\mathbb{M})$ such that:

- for any $z\in\mathbb{M}$ ν-a.s., $s\longmapsto\gamma(s,z)\in\mathcal{P}(\mathcal{F})$;
- $\forall(s,u)\in\mathbb{R}^+\times\mathbb{M},\ \gamma(s,u)\in\mathbb{M}$.

For any $s\in\mathbb{R}^+$, we say that γ is μ-invertible with inverse denoted $\gamma^*:\mathfrak{M}\times\mathbb{R}^+\times\mathbb{M}\to\mathbb{M}$ if given $z\in\mathbb{M}$, for almost all $\omega\in\mathfrak{M}$, $s\in\mathbb{R}^+$:

$$\gamma(\omega,s,\gamma^*(\omega,s,z)) = \gamma^*(\omega,s,\gamma(\omega,s,z)) = z \text{ for all } s\in\mathbb{R}^+ \quad \mathrm{d}t\otimes\mu\text{-a.s.}$$

The transformation of mark is a very common operation on Poisson measures. It has been an important development in the theory of Malliavin calculus for perturbation analysis (see Bichteler and Jacod (1983)).

Remark 4.3.2. The random transformation of mark is a particular case of random transformation of marked point processes that shifts each point of a configuration $\omega\in\mathfrak{N}$ in some direction τ.

Definition 4.3.3. Let X a random process on $\mathbb{R}^+\times\mathbb{M}$ and γ invertible transformation of mark. Then, the changed random process is defined by

$$\mathbf{\Gamma}(X)(t,A) = \int_0^t\int \mathbf{1}_A(\gamma(s,z))\,X(\,\mathrm{d}s,\,\mathrm{d}z) \quad \text{for } (t,A)\in\mathbb{R}^+\times\mathcal{B}(\mathbb{M}).$$

If $X\in\mathfrak{M}$, $\mathbf{\Gamma}(X)\in\mathfrak{M}$.

Lemma 4.3.4 *Let γ an invertible transformation, we have:*

$$\mathcal{F}_t^{\mathbf{\Gamma}(X)}\vee\sigma(\gamma^*(s,z),\,s\le t,z\in\mathbb{M}) = \mathcal{F}_t^X \tag{4.3.1}$$

Proof. We have by construction of $\mathbf{\Gamma}(X)$, $\mathcal{F}_t^{\mathbf{\Gamma}(X)} = \sigma(\mathbf{\Gamma}(X)(s,A),\,s\le t,A\in\mathcal{B}(\mathbb{M}))\subset\mathcal{F}_t^X$

relative to the Poisson measure X. By predictability of γ^*, $\mathcal{F}_t^{\mathbf{\Gamma}(X)} \vee \sigma(\gamma^*(s,z),\, s \le t, z \in \mathbb{M}) \subset \mathcal{F}_t^X$. Conversely, for $A \in \mathcal{F}_t^N$, there exists a sequence of measurable functions $(\chi_q)_{q\in\mathbb{N}}$ such that for any q, χ_q is measurable function from $(\mathbb{R}^+ \times \mathbb{M})^q$ to $\{0,1\}$:

$$\begin{aligned}\mathbf{1}_A &= \sum_{q=0}^{\infty} \chi_q(T_1(X),\dots,T_q(X),Z_1(X),\dots,Z_q(N))\mathbf{1}_{\{T_q(X)<t\le T_{q+1}(X)\}}\\ &= \sum_{q=0}^{\infty} \chi_q^{\mathbf{\Gamma}(X)}\mathbf{1}_{\{T_q(\mathbf{\Gamma}(X))<t\le T_{q+1}(\mathbf{\Gamma}(X))\}}\end{aligned}$$

where

$$\begin{aligned}\chi_q^{\mathbf{\Gamma}(X)} = \chi_q(T_1(\mathbf{\Gamma}(X)),\dots,T_q(\mathbf{\Gamma}(X)),\gamma^*(X,T_1(\mathbf{\Gamma}(X)),Z_1(\mathbf{\Gamma}(X))),\\ \dots,\gamma^*(X,T_q(\mathbf{\Gamma}(X)),Z_q(\mathbf{\Gamma}(X)))).\end{aligned}$$

Because $\gamma^*(N,\cdot)$ is measurable from $\mathbb{R}^+ \times \mathbb{M}$ to $\mathbb{R}^+$, it is a limit of simple functions:

$$\begin{aligned}&\gamma^*(X,T_k(\mathbf{\Gamma}(X)),Z_k(\mathbf{\Gamma}(X)))\\ &\qquad= \lim_{n\to\infty}\sum_{i=0}^{2^n-1}\sum_{j=0}^{2^m-1}\gamma^*(X,\frac{i}{2^n},\frac{j}{2^m})\mathbf{1}_{[\frac{(i-1)}{2^n},\frac{i}{2^n})}(T_k(\mathbf{\Gamma}(X)))\mathbf{1}_{[\frac{(j-1)}{2^m},\frac{j}{2^m})}Z_k(\mathbf{\Gamma}(X)).\end{aligned}$$

Hence the result. □

In the remainder, we consider as canonical marked point processes which is a stationary marked Poisson point process N with intensity measure $\mathrm{d}s\otimes\nu(\mathrm{d}z)$ for a σ-finite measure ν on $(\mathbb{M},\mathcal{B}(\mathbb{M}))$ where $\mathrm{d}s$ is the Lebesgue measure on $\mathbb{R}$ and the associated probability measure π on $(\mathfrak{M},\mathcal{F}_\infty)$ by Theorem 4.1.7. We denote Γ the process $\mathbf{\Gamma}(N)$. Under π, we identify ω and the canonical process N. We denote by π_σ the unique probability measure on $(\mathfrak{M},\mathcal{F}_\infty^N)$ such that $N^p = \sigma(s,z)\nu(\mathrm{d}z)\,\mathrm{d}s$ for σ a random process on $\mathfrak{M}$. When $\sigma=1$, we recover $\pi=\pi_{\mathrm{Id}}$. We keep the notation π for sake of simplicity.

4.3.2 Change of measures and Girsanov theorem

Here we consider a probability π' with the basic hypothesis that π' is locally absolutely continuous with respect to π along $\mathcal{F}$. We write it $\pi' \ll_{\mathrm{loc},\mathcal{F}} \pi$. Our aim is to compute the dual predictable projection of N under π'. This is the object of the so-called Girsanov theorems. The next theorem is a combination of (Decreusefond, 1998, p.500) and (Jacod, 1975, theorem 4.5). The proof of the criterion for the Radon-Nikodym derivative to be non-zero can found in the comprehensive book (Jacod and Shiryaev, 2003, Theorem 5.19 p.195 and Corollary 5.22 p.195).

Theorem 4.3.5 *Consider a filtration $\mathcal{F}=(\mathcal{F}_t,\, t\ge 0)$ on $\mathfrak{M}$, and assume that under π, the process*

$$\tilde{N}\,:\,t\longmapsto N\big(t,A\big)-\int_0^t\int \mathbf{1}_A(z)\nu(\mathrm{d}z)\,\mathrm{d}s$$

is a $\mathcal{F}$-local martingale. If a probability measure π' on $(\mathfrak{M},\mathcal{F}_\infty^N)$ is locally absolutely *continuous with respect to π along $\mathcal{F}$ then there exists a unique process ϕ on $\mathbb{R}^+ \times \mathbb{M}$ which is non-negative π-a.s. and $\mathcal{F}$-predictable such that*

$$\forall t \geq 0,\ \pi'\left(\int_0^t \int_{\mathbb{M}} \left(1 - \sqrt{\phi(s,z)}\right)^2 \nu(\mathrm{d}z)\ \mathrm{d}s < \infty\right) = 1 \tag{4.3.2}$$

and for any $A \in \mathcal{B}(\mathbb{M})$,

$$t \longmapsto N\big(t, A\big) - \int_0^t \int_A \phi(s,z)\nu(\mathrm{d}z)\ \mathrm{d}s \text{ is a } (\mathcal{F}, \pi_\phi)\text{-local martingale.} \tag{4.3.3}$$

We refer to ϕ as the Girsanov factor of $\pi' := \pi_\phi$ with respect to π. Moreover, assume (A1), for any integer $m \geq 1$, let

$$S_m = \inf\left\{t \in \mathbb{R}^+,\ \int_0^t \int_{\mathbb{M}} \left(1 - \sqrt{\phi(s,z)}\right)^2 \nu(\mathrm{d}z)\ \mathrm{d}s \geq m\right\}.$$

Then, with the previously introduced notations,

$$\begin{aligned}\Lambda_\phi(t) :&= \left.\frac{\mathrm{d}\pi_\phi}{\mathrm{d}\pi}\right|_{\mathcal{F}_t} \\ &= \begin{cases} \mathcal{E}\left(\delta_{\tilde{N}}\phi\right)(t) & \text{if } t \leq S_m, \\ 0 & \text{if } t \geq \limsup_m S_m. \end{cases}\end{aligned}$$

If π_ϕ is absolutely continuous with respect to π on $\mathcal{F}_\infty$ then

$$\pi_\phi\left(\int_0^\infty \int \left(1 - \sqrt{\phi(s,z)}\right)^2 \nu(\mathrm{d}z)\,\mathrm{d}s < \infty\right) = 1. \tag{4.3.4}$$

For the converse part, let us suppose (A1). Consider ϕ a non negative $\mathcal{F}$-predictable process and π_ϕ the probability measure on $(\mathfrak{M}, \mathcal{F}^N)$, which satisfies (4.3.2) and (4.3.3). Then, π_ϕ is locally absolutely continuous with respect to π along $\mathcal{F}$.
Finally, the probability measure π_ϕ is absolutely continuous with respect to π on $(\mathfrak{M}, \mathcal{F}_\infty)$ if and only if (4.3.4) is satisfied.

Definition 4.3.6. Let $\mathcal{F}$ a filtration relative to N, we introduce the following subset of $\mathbb{F}^1(\mathcal{P}(\mathcal{F}), \nu \otimes \mathrm{d}s)$ denoted $\mathbb{F}^1_2(\mathcal{P}(\mathcal{F}), \nu \otimes \mathrm{d}s; \pi)$ such that the following properties hold for an element ϕ:

1. $\pi_\phi \ll_{\text{loc},\mathcal{F}} \pi$
2. $\pi_\phi\left(\int_0^\infty \int \left(1 - \sqrt{\phi(s,z)}\right)^2 \nu(\mathrm{d}z)\,\mathrm{d}s < \infty\right) = 1.$

We recall from Coutin and Decreusefond (2023) some necessary and sufficient conditions for the equivalence of π and π_ϕ, see also Proposition (7.11) of Jacod (1979).

Lemma 4.3.7 *Let μ' a measure absolutely continuous with respect to μ on $\mathcal{F}_\infty$ and set*

$$\Lambda(t) = \left.\frac{\mathrm{d}\mu'}{\mathrm{d}\mu}\right|_{\mathcal{F}_t}. \tag{4.3.5}$$

Then μ' and μ are equivalent if and only the following two conditions are satisfied:

i) The local martingale $(\Lambda(t),\, t \ge 0)$ is uniformly integrable, i.e. there exists $\Lambda \in L^1(\mathfrak{M} \times \mathbb{R} \to \mathbb{R}; \mu)$ such that

$$\Lambda(t) = \mathbf{E}_\mu \left[\Lambda \,|\, \mathcal{F}_t\right].$$

ii) The random variable Λ is positive μ-a.s.

We fix $\mathbb{M} = \mathbb{R}^+$ endowed with the Lebesgue measure. The central notion introduced by Bismut is the Girsanov transform which is key to our investigations.

Definition 4.3.8 (Girsanov transform)**.** Let $\phi \in \mathcal{P}^{++}(\mathcal{F})$. Define γ_ϕ^ρ, for $\pi \otimes \mathrm{d}s$ all $\omega \in \mathfrak{M}$ and $u \ge 0$ as a Lebesgue integral:

$$\gamma_\phi^\rho(\omega, s, z) = \int_0^z \phi(\omega, s, u)\rho(u)\ \mathrm{d}u.$$

Then, setting $r(z) = \int_0^z \rho(u)\ \mathrm{d}u$, the random transformation of mark

$$\gamma_\phi(\omega, s, \cdot) = r^{-1} \circ \gamma_\phi^\rho(\omega, s, \cdot).$$

is a C^1-diffeomorphism and called *Girsanov transform.* Moreover, it is π-invertible in the sense of Definition 4.3.1 with inverse denoted γ_ϕ^*. In the same vein, we define the process:

$$\phi^\dagger(\omega, s, z) = \frac{1}{\phi(\omega, s, \gamma_\phi^*(\omega, s, z))}.$$

Using the Definition 4.3.3, given ϕ, we have the triplet $(\gamma_\phi, \mathbf{\Gamma}_\phi, \phi^\dagger)$ that characterizes the Girsanov transform.

According to Fujisaki Fujisaki and Komatsu (2021), the Girsanov transform refers to absolutely continuous transform of laws of stochastic processes. There is no formal definition in Bichteler et al. (1987) which dedicates a paragraph to it.

Remark 4.3.9. We have for $s \in \mathbb{R}^+$:

$$(\gamma_\phi(s, \cdot))'(u) = \frac{\phi(s, u)\rho(u)}{\rho(\gamma_\phi(s, u))}.$$

and

$$(\gamma_\phi^*(s, \cdot))'(u) = \frac{\rho(u)}{\rho(\gamma_\phi^*(s, u))\phi(s, \gamma_\phi^*(s, u))} = \frac{\rho(u)\phi^\dagger(s, u)}{\rho(\gamma_\phi^*(s, u))}. \tag{4.3.6}$$

For instance, if $\rho = 1$, we have:

$$(\gamma_\phi(s,\cdot))'(z) = \phi(s,z) \text{ and } (\gamma^*_\phi(s,\cdot))'(z) = \phi^\dagger(s,z).$$

Remark 4.3.10. The generalization of the Girsanov transform to $\mathbb{R}^d$ for $d \geq 2$ is immediate as the remark 4.3.9 still holds. The difference with the case $d = 1$ lies in the stochastic integration rules. Because $\phi > 0$, π-a.s., it corresponds for any $\omega \in \mathfrak{M}$ π-a.s. and each $s \in \mathbb{R}^+$ to:

$$\gamma'_\phi(\omega,s,x) = |\det(D\gamma_\phi(\omega,s,\cdot))(x)|$$

for U an open set in $\mathbb{R}^d$ and $\gamma^*_\phi(s,\cdot) : U \to \mathbb{R}^d$ an injective differentiable function with continuous partial derivatives, the Jacobian of which is nonzero for $x \in U$. We notice that not all the components of γ^*_ϕ need to be nonnegative for $d \geq 2$. For sake of readability, this generalization is omitted.

Denote better the multi-dimensional derivative

Lemma 4.3.11 *For $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \nu \otimes \mathrm{d}t)$, then we have:*

$$\mathbf{\Gamma}_\phi\left(\int_0^\cdot \int f(N,\, s, z)\nu(\,\mathrm{d}z)\,\mathrm{d}s\right) = \int_0^\cdot \int f(N,\, s, \gamma^*_\phi(N,s,z))\phi^\dagger(N,s,z)\nu(\,\mathrm{d}z)\,\mathrm{d}s. \quad (4.3.7)$$

and:

$$\mathbf{\Gamma}_\phi\left(\int_0^\cdot \int f(N,\, s, z)N(\,\mathrm{d}s,\,\mathrm{d}z)\right) = \int_0^\cdot \int f(N,\, s, \gamma^*_\phi(N,s,z))\,\mathrm{d}\Gamma_\phi(s,z). \quad (4.3.8)$$

Proof. For $t \in \mathbb{R}^+$ and $A \in \mathcal{B}(\mathbb{R}^+)$,

$$\mathbf{\Gamma}_\phi\left(\int_0^\cdot \int f(N,\, s, z)\nu(\,\mathrm{d}z)\,\mathrm{d}s\right)(N,t,A) = \int_0^\cdot \int \mathbf{1}_A(\gamma_\phi(N,s,z))f(N,s,z)\nu(\,\mathrm{d}z)\,\mathrm{d}s.$$

In particular, for $a, b \in \mathbb{R}^+$ with $a < b$, $C \in \mathcal{B}(\mathbb{M})$ and $B \in \mathcal{F}_a$, let

$$f(N,s,z) = B(N)\mathbf{1}_{(a,b]}(s)\mathbf{1}_C(z).$$

We have by the theorem of change of variables in terms of Lebesgue measure on $\mathbb{R}$, with $u = \gamma_\phi(N,s,z)$ for each $s \in [0,t]$:

$$\int \mathbf{1}_A(\gamma_\phi(N,s,z))\mathbf{1}_C(z)\nu(\,\mathrm{d}z) = \int \mathbf{1}_A(u)\mathbf{1}_C(\gamma^*_\phi(N,s,u))\phi^\dagger(N,s,u)\nu(\,\mathrm{d}u)\,\mathrm{d}s$$

since $(\gamma^*_\phi)'(N,s,\cdot)(u)\rho(\gamma^*_\phi(N,s,u)) = \frac{1}{\phi(N,s,\gamma^*_\phi(N,s,u))}\rho(u)$. Then,

$$\int \mathbf{1}_A(\gamma_\phi(N,s,z))f(N,\, s, \gamma^*_\phi(s,u))\nu(\,\mathrm{d}z)\,\mathrm{d}s = \int \mathbf{1}_A(u)f(N,\, s, \gamma^*_\phi(s,u))\phi^\dagger(N,s,u)\nu(\,\mathrm{d}u)\,\mathrm{d}s.$$

By density of simple processes in $\mathbb{F}^1(\mathcal{F}^N, \nu \otimes \mathrm{d}s)$, this yields (4.3.7). We apply the same change of variable for the proof of the second identity.

$$\begin{aligned}\mathbf{\Gamma}_\phi\left(\int_0^{\cdot}\int f(N,\, s, z)N(\,\mathrm{d}s,\,\mathrm{d}z)\right)(t, A) &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(N, s, z))f(N,\, s, z)N(\,\mathrm{d}s,\,\mathrm{d}z)\\ &= \int_0^t \int \mathbf{1}_A(u)f(N,\, s, \gamma^*_\phi(N, s, u))\Gamma(\,\mathrm{d}s,\,\mathrm{d}u).\end{aligned}$$

Since it holds for each $A \in \mathcal{B}(\mathbb{R}^+)$, it yields (4.3.8). □

Corollary 4.3.12 *Let $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$, then $\phi^\dagger \in \mathcal{P}^{++}(\mathcal{F}^{\Gamma_\phi})$.*

Remark 4.3.13. The factor $\phi^\dagger$ is a Girsanov factor as well as ϕ.

4.3.3 Main theorem

Theorem 4.3.14 — Girsanov theorem for Poisson measures. *Let $\pi' \ll_{loc,\mathcal{F}^N} \pi$ and $\phi \in \mathcal{P}(\mathcal{F}^N)$ denote its Girsanov factor, i.e. $\pi' = \pi_\phi$. Assume that ϕ belongs to $\mathbb{F}^1_2(\mathcal{P}^{++}(\mathcal{F}^N), \nu\otimes \mathrm{d}t, \pi_\phi)$. Then, with our previous notations, the distribution of the process Γ_ϕ under π_ϕ is the distribution of N under π. This means that for any bounded measurable $f : \mathfrak{M} \to \mathbb{R}$, for any $t \in \mathbb{R}^+$,*

$$\mathbf{E}_\pi\left[f(\Gamma^t_\phi)\ \Lambda_\phi(t)\right] = \mathbf{E}_\pi\left[f(N^t)\right], \tag{4.3.9}$$

where X^t is the process X stopped at time t.

Proof. The Radon-Nikodym derivative is given by Theorem 4.3.5.

$$\begin{aligned}\Lambda_\phi(t) &= \left.\frac{\mathrm{d}\pi_\phi}{\mathrm{d}\pi}\right|_{\mathcal{F}^N_t}\\ &= \exp\left(\int_0^t \int \log\left(\phi(s, z)\right)\,\mathrm{d}N(s, z) - \int_0^t\int \Big(\phi(s, z) - 1\Big)\nu(\mathrm{d}z)\,\mathrm{d}s\right)\end{aligned} \tag{4.3.10}$$

We recall that for each $A \in \mathcal{B}(\mathbb{M})$, the compensator of $\Gamma_\phi(\cdot, A)$ under π is $\int_0^{\cdot} \mathbf{1}_A(\gamma_\phi(s, z))\nu(\,\mathrm{d}z)$. Thus by Lemma 4.3.11 and Remark 4.3.9,

$$\begin{aligned}R(t) &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(s, z))\phi(s, z)\big(N(\,\mathrm{d}s,\,\mathrm{d}z) - \nu(\,\mathrm{d}z)\,\mathrm{d}s\big)\\ &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(s, z))\phi(s, z)N(\,\mathrm{d}s,\,\mathrm{d}z) - t\int \mathbf{1}_A(z)\nu(\,\mathrm{d}z)\end{aligned}$$

is a $(\mathcal{F}^N, \pi)$ local martingale. The standard Girsanov's theorem for local martingales (Schuppen and Wong, 1974, Theorem 3.2.) says that

$$R(t) - \int_0^t \frac{1}{\Lambda_\phi(s)}\,\mathrm{d}[R, \Lambda_\phi](s)$$

is a $(\mathcal{F}^N, \pi_\phi)$- local martingale. Note that R and Λ_ϕ have the same jump times as Γ_ϕ, hence

$$\begin{aligned}\int_0^t \frac{1}{\Lambda_\phi(s)} \, \mathrm{d}[R, \Lambda_\phi](s) &= \sum_{s \le t, \Delta N(s) \neq 0} \frac{1}{\Lambda_\phi(s)} \Delta R(s)\, \Delta\Lambda_\phi(s) \\ &= \sum_{s \le t, \Delta N(s) \neq 0} \left(1 - \frac{\Lambda_\phi(s^-)}{\Lambda_\phi(s)}\right) \Delta R(s) \\ &= \sum_{T_n \le t,\, Z_n \neq 0} (1 - \frac{1}{\phi(T_n, Z_n)}) \mathbf{1}_A(\gamma_\phi(T_n, Z_n))\phi(T_n, Z_n) \cdot \\ &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(s,z))\phi(s,z) \, \mathrm{d}N(s,z) - \Gamma_\phi(t, A).\end{aligned}$$

Thus, we have

$$R(t) - \int_0^t \frac{1}{\Lambda_\phi(s)} \, \mathrm{d}[R, \Lambda_\phi](s) = \Gamma_\phi(t, A) - t \int \mathbf{1}_A(z) \, \mathrm{d}z.$$

This means that the random measure associated to Γ_ϕ has $(\mathcal{F}^N, \pi_\phi)$-dual predictable projection $\nu(\mathrm{d}z) \otimes \mathrm{d}s$. According to Theorem 4.1.2 by the characterization of homogeneous Poisson measure, the random measure is an $\mathcal{F}^N$-adapted homogeneous Poisson measure of control measure $\nu(\,\mathrm{d}z)\,\mathrm{d}s$ under π_ϕ. □

The version of the quasi-invariance theorem that we use subsequently is the following.

Theorem 4.3.15 — Quasi-invariance. *Let $\pi_\phi \ll_{loc, \mathcal{F}^\Gamma} \pi$. Then the distribution of Γ_ϕ under π_ϕ is the distribution of N under π, i.e.:*

$$\mathbf{E}_\pi \left[f(\Gamma_\phi^t)\; \Lambda^\dagger_{\phi^\dagger}(t) \right] = \mathbf{E}_\pi \left[f(N^t) \right], \tag{4.3.11}$$

with

$$\Lambda^\dagger_{\phi^\dagger}(t) = \frac{\mathrm{d}\pi_\phi}{\mathrm{d}\pi}\Big|_{\mathcal{F}^\Gamma_t} \tag{4.3.12}$$

$$= \exp\left(- \int_0^t \int \log\left(\phi^\dagger(s,u)\right) \, \mathrm{d}\Gamma_\phi(s,u) + \int_0^t \int \left(\phi^\dagger(s,u) - 1\right) \nu(\mathrm{d}u) \, \mathrm{d}s \right) \tag{4.3.13}$$

Proof. The Girsanov theorem yields for any bounded measurable $f : \mathfrak{M} \to \mathbb{R}$, for any $t \in \mathbb{R}^+$:

$$\mathbf{E}_{\pi_\phi} \left[f(\mathbf{\Gamma}_\phi)\; \Lambda_\phi(t) \right] = \mathbf{E}_\pi \left[f \right].$$

We proceed with the change of variable $z = \gamma_\phi(s,u)$ over $\mathbb{R}^+$. By Lemma 4.3.11,

$$\begin{aligned}\log \Lambda_\phi(t) &= \int_0^t \int \log\Big(\phi(s, \gamma^*_\phi(s,u))\Big) \, \mathrm{d}\Gamma_\phi(s,u) - \int_0^t \int \Big(\phi(s, \gamma^*_\phi(s,u)) - 1\Big) \phi^\dagger(s,u) \nu(\,\mathrm{d}u) \, \mathrm{d}s \\ &= - \int_0^t \int \log\left(\phi^\dagger(s,u)\right) \, \mathrm{d}\Gamma_\phi(s,u) + \int_0^t \int \left(\phi^\dagger(s,u) - 1\right) \nu(\mathrm{d}u) \, \mathrm{d}s.\end{aligned}$$

By Lemma 4.3.12, $\Lambda^\dagger_{\phi^\dagger}(t)$ is $\mathcal{F}_t^{\Gamma_\phi}$-measurable. The proof is thus complete. □

We now prove that well behaved mark changes induce locally absolutely continuous probability on $\mathfrak{M}$.

Theorem 4.3.16 *Let ϕ belong to $\mathcal{P}^{++}(\mathcal{F}^N)$ such that π_ϕ is equivalent to π on $\mathcal{F}_\infty$. Then $\mathbf{\Gamma}_{\phi\#}\pi$ is equivalent to π on $\mathcal{F}_\infty$.*

Proof. According to Lemma 4.3.7, the martingale

$$(\Lambda_\phi(t),\, t \geq 0)$$

is uniformly integrable and we can let t go to infinity in (4.3.9) to obtain

$$\mathbf{\Gamma}_{\phi\#}\pi_\phi = \pi.$$

Since π_ϕ is equivalent to π on $\mathcal{F}_\infty$, $\mathbf{\Gamma}_{\phi\#}\pi_\phi$ is equivalent to $\mathbf{\Gamma}_{\phi\#}\pi$. As a consequence,

$$\pi = \mathbf{\Gamma}_{\phi\#}\pi_\phi \sim \mathbf{\Gamma}_{\phi\#}\pi.$$

The proof is thus complete. □

We conclude this section by introducing a sufficient condition such that the equivalence between probability measures holds.

Definition 4.3.17. The most restricted class we consider is $\mathbb{F}_\infty(\mathcal{P}^{++}(\mathcal{F}), \pi_\sigma)$ of processes $\varphi \in \mathcal{P}^{++}(\mathcal{F})$ for which there exist $c \in (0,1)$ such that

$$c \leq \varphi(s,z) \leq \frac{1}{c},\ \forall s \geq 0,\ z \in \mathbb{M}\ \pi_\sigma - \text{a.s.}$$

Remark 4.3.18. Note that if ϕ belongs to $\mathbb{F}_\infty(\mathcal{P}^{++}(\mathcal{F}^N), \pi_\sigma)$ then there exists $C < +\infty$ such that $\pi -$ a.s., $0 < \Lambda_\phi(N,t) \leq C$. Hence, $(\Lambda_\phi(N,t),\, t \geq 0)$ is uniformly integrable.

4.4 Invertibility

Those definitions are analog to the ones of Üstünel (2014) relative to the Wiener space. We add the absolute continuity condition as we consider Girsanov transforms as absolute continuous transformations.

Definition 4.4.1. For a probability μ on $(\mathfrak{M}, \mathcal{F}_\infty)$, a map $\mathfrak{Y}\ : \mathfrak{M} \to \mathfrak{M}$ is μ-left invertible if and only if $\mathfrak{Y}_{\#}\mu \ll \mu$ along $\mathcal{F}_\infty$ and there exists $\mathfrak{Z} : \mathfrak{M} \to \mathfrak{M}$ such that $\mathfrak{Z} \circ \mathfrak{Y} = \text{Id}_{\mathfrak{N}}$, μ-a.s.
The map $\mathfrak{Y}\ : \mathfrak{M} \to \mathfrak{M}$ is μ-right invertible if and only if there exists $\mathfrak{Z} : \mathfrak{M} \to \mathfrak{M}$ such that $\mathfrak{Z}_{\#}\mu \ll \mu$ along $\mathcal{F}_\infty$ and $\mathfrak{Y} \circ \mathfrak{Z} = \text{Id}_{\mathfrak{N}}$, μ-a.s.
The map $\mathfrak{Y}$ is μ-invertible if it is both μ-left and μ-right invertible.

Because the invertibility of sample path remains with respect to π, we omit to mention it.

Lemma 4.4.2 *If there exists $\mathfrak{Z}$ such that $\mathfrak{Z} \circ \mathfrak{Y} = \text{Id}_{\mathfrak{M}}$, π-a.s. then $\mathfrak{Y} \circ \mathfrak{Z} = \text{Id}_{\mathfrak{M}}$, $\mathfrak{Y}_{\#}\pi$-a.s. If additionally, $\mathfrak{Y}_{\#}\pi$ is equivalent to π and $\mathfrak{Z}_{\#}\pi \ll \pi$, then $\mathfrak{Y}$ is invertible and $\mathfrak{Z}_{\#}\pi$ is equivalent to π.*

Proof. We have

$$\begin{aligned}\mathfrak{Y}_{\#}\pi\Big(\mathfrak{Y}\circ\mathfrak{Z}=\mathrm{Id}_{\mathfrak{M}}\Big)&=\pi\Big(\mathfrak{Y}\circ\mathfrak{Z}\circ\mathfrak{Y}=\mathfrak{Y}\Big)\\&=\pi\Big(\mathfrak{Y}=\mathfrak{Y}\Big)\\&=1.\end{aligned}$$

The first assertion follows.
If the two measures $\mathfrak{Y}_{\#}\pi$ and π are equivalent, then $\mathfrak{Y}\circ\mathfrak{Z}=\mathrm{Id}_{\mathfrak{M}}$ π-almost-surely thus $\mathfrak{Y}$ is invertible. Let A such that $\mathfrak{Z}_{\#}\pi(A)=0$. This means

$$\mathbf{E}_\pi\left[\mathbf{1}_A\circ\mathfrak{Z}\right]=0.$$

Since $\mathfrak{Y}_{\#}\pi$ is equivalent to π, we get

$$0=\mathbf{E}_\pi\left[\mathbf{1}_A\circ\mathfrak{Z}\circ\mathfrak{Y}\right]=\mathbf{E}_\pi\left[\mathbf{1}_A\right]$$

hence $\pi\ll\mathfrak{Z}_{\#}\pi$ and the equivalence follows. □

We begin by a technical lemma which states the composition by the change of marks $\mathbf{\Gamma}_\phi$.

Lemma 4.4.3 *For $\psi\in\mathcal{P}^{++}(\mathcal{F}^N,\pi)$, we define the Girsanov transform γ_ψ according to Definition 4.3.8. Then, we have*

$$(\mathbf{\Gamma}_\psi\circ\mathbf{\Gamma}_\phi(N))(t,A)=\int\int\mathbf{1}_{[0,t]}(s)\mathbf{1}_A(\gamma_\psi(\Gamma_\phi,s,\gamma_\phi(N,s,z)))\ \mathrm{d}N(s,z).\tag{4.4.1}$$

Moreover, for $f\in\mathbb{F}^1(\mathcal{P}(\mathcal{F}),\nu\otimes\mathrm{d}t)$,

$$\left(\int_0^{\cdot}\int f(N,\,s,\,z)\nu(\,\mathrm{d}z)\ \mathrm{d}s\right)\circ\mathbf{\Gamma}_\phi=\int_0^{\cdot}\int f\big(\Gamma_\phi,\,s,\,z\big)\nu(\,\mathrm{d}z)\ \mathrm{d}s.\tag{4.4.2}$$

and

$$\left(\int_0^{\cdot}\int f(N,\,s\,,z)\ \mathrm{d}N(s,z)\right)\circ\mathbf{\Gamma}_\phi=\int_0^{\cdot}\int f(\Gamma_\phi,\,s,\,z)\ \mathrm{d}\Gamma_\phi(s,z).\tag{4.4.3}$$

Proof. For $t\geq 0$ and $A\in\mathcal{B}(\mathbb{M})$, we have by definition:

$$(\mathbf{\Gamma}_\psi\circ\mathbf{\Gamma}_\phi)(X)\,(t,A)=\int\int\mathbf{1}_{[0,t]}(s)\mathbf{1}_A(\gamma_\psi(\mathbf{\Gamma}_\phi(X),s,z))\,\kappa_\gamma(X)(\,\mathrm{d}s,\,\mathrm{d}z)\tag{4.4.4}$$

$$=\int\int\mathbf{1}_{[0,t]}(s)\mathbf{1}_A(\gamma_\psi(\mathbf{\Gamma}_\phi(X),s,\gamma_\phi(X,s,z)))\ \mathrm{d}X(s,z).\tag{4.4.5}$$

Because of the density of simple processes in $\mathbb{F}^1(\mathcal{F}^N,\nu\otimes\mathrm{d}s)$, we can apply the composition rule (4.4.4) to processes

$$(t,A)\mapsto\int_0^t\int\mathbf{1}_A(z)f(N,s,z)\ \mathrm{d}N^p(s,z)$$

and

$$(t, A) \mapsto \int_0^t \int \mathbf{1}_A(z) f(N, s, z) \, \mathrm{d}N(s, z)$$

evaluated at $(t, \mathbb{M})$ as to obtain (4.4.2) and (4.4.3). □

Corollary 4.4.4 *For $\psi \in \mathcal{P}^{++}(\mathcal{F}^N, \pi)$, $\mathbf{\Gamma}_\psi$ is the π-left inverse of $\mathbf{\Gamma}_\phi$ if and only if $\mathbf{\Gamma}_{\phi\#}\pi \ll \pi$ and:*

$$\gamma_\psi(\Gamma_\psi, s, \gamma_\phi(N, s, z)) = z \ \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \ \pi\text{-a.s.} \tag{4.4.6}$$

or equivalently either of those equations:

- $$\gamma_\psi^*(\Gamma_\phi, s, z) = \gamma_\phi(N, s, z) \ \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \ \pi\text{-a.s.} \tag{4.4.7}$$

- $$\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)) \times \phi(N, s, z) = 1 \ \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \ \pi\text{-a.s..} \tag{4.4.8}$$

Proof. We have:

$$\mathbf{\Gamma}_\psi \circ \mathbf{\Gamma}_\phi = \mathrm{Id}_{\mathfrak{M}} \ \pi\text{-a.s.} \iff \gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)) = z \ \pi \otimes \mathrm{d}s \otimes \mathrm{d}z\text{-a.s.} \tag{4.4.9}$$

Since γ_ψ is π-invertible, it is equivalent to

$$\gamma_\psi^*(\Gamma_\phi, s, z) = \gamma_\phi(N, s, z) \ \pi \otimes \mathrm{d}s \otimes \mathrm{d}z\text{-a.s..}$$

Let us show that it is also equivalent to (4.4.8). By differentiation of (4.4.6), we obtain

$$\big(\gamma_\psi(\Gamma_\phi, s, \cdot)\big)'\big(\gamma_\phi(N, s, z)\big) \times (\gamma_\phi(N, s, \cdot))'(z) = 1, \ \pi - \text{a.s.}$$

Using Remark 4.3.9, we obtain

$$\begin{aligned}\big(\gamma_\psi(\Gamma_\phi, s, \cdot)\big)'\big(\gamma_\phi(N, s, z)\big) \times (\gamma_\phi^*(N, s, \cdot))'(z) \\ &= \frac{\rho(\gamma_\phi(N, s, z))\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))}{\rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)))} \times \frac{\rho(z)\phi(N, s, z)}{\rho(\gamma_\phi(N, s, z))} \\ &= \psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))\phi(N, s, z) = 1\end{aligned}$$

Conversely, assume there exists $\psi \in \mathcal{P}^{++}(\mathcal{F}^N)$ such that

$$\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))\phi(N, s, z) = 1 \ \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \ \pi\text{-a.s..}$$

Given γ_ψ defined according to Definition 4.3.8, it yields:

$$\begin{aligned}\big(\gamma_\psi(\Gamma_\phi, s, \cdot)\big)'\big(\gamma_\phi(N, s, z)\big) \times (\gamma_\phi(N, s, \cdot))'(z) &= \frac{\rho(\gamma_\phi(N, s, z))\rho(z)}{\rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)))\rho(\gamma_\phi(N, s, z))} \\ &= \frac{\rho(z)}{\rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)))}.\end{aligned}$$

Hence,

$$\gamma_\psi(\Gamma_\phi, s, \cdot)'(\gamma_\phi(N, s, \cdot)) \times \rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))) = \rho(z)$$
$$(r \circ \gamma_\psi(\Gamma_\phi(N), s, \gamma_\phi^*(N, s, \cdot)))' = r'.$$

By integration,

$$r \circ \gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, \cdot)) = r.$$

Because r is a bijection from $\mathbb{M}$ onto itself, it follows that

$$\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, \cdot)) = \mathrm{Id}_{\mathbb{M}}\,.$$

□

By definition, $\mathcal{F}^N$ is smaller than $\mathcal{F}^\Gamma$. If we assume left invertibility of the map $\mathbf{\Gamma}$, we have a stronger result.

Theorem 4.4.5 *Let $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$. Then, $\mathcal{F}^\Gamma = \mathcal{F}^N$ if and only if $\mathbf{\Gamma}_\phi$ admits a left inverse.*

Proof. If $\mathbf{\Gamma}_\phi$ is left invertible then there exists γ_ψ such that (4.4.7) holds:

$$\gamma_\psi(\mathbf{\Gamma}(\omega), s, z) = \gamma_\phi^*(\omega, s, z)\ \pi \otimes \mathrm{d}s \otimes \nu\text{-a.s.}$$

This means that:

$$\gamma_\phi^*(N, t, \cdot) \text{ is } \mathcal{F}_t^\Gamma\text{-measurable.}$$

Hence, by lemma 4.3.4 $\mathcal{F}^\Gamma = \mathcal{F}^N$.
Conversely, if, for all $t \in \mathbb{R}^+$, $\mathcal{F}_t^N = \mathcal{F}_t^\Gamma$, then for all $t \in \mathbb{R}^+$ there exists a sequence of predictable processes $(\bar{\gamma}_\psi^t)_{t\in\mathbb{R}^+}$ such that:

$$\bar{\gamma}_\psi^t(\Gamma, z) = \gamma_\phi^*(N, t, z),\ \ \pi \otimes \mathrm{d}z - \text{a.s.}$$

Denoting $\bar{\gamma}_\psi(\mathbf{\Gamma}(\omega), t, z) = \bar{\gamma}_\psi^t(\Gamma, z)$ for all $t \in \mathbb{R}^+$, there exists a full probability set B such that

$$\bar{\gamma}_\psi(\mathbf{\Gamma}(\omega), t, z) = \gamma_\phi^*(\omega, t, z), \forall t \in \mathbf{Q},\ \ \forall z \in \mathbf{Q},\ \ \forall \omega \in B. \tag{4.4.10}$$

Let

$$\gamma_\psi(\Gamma, t, z) = \begin{cases} \gamma_\phi^*(N, t, z) & \text{if } t \in \mathbf{Q} \\ \lim_{r_n \to t, r_n \in \mathbf{Q}\cap[0,t]} \gamma_\phi^*(N, t, z) & \text{if } t \notin \mathbf{Q}. \end{cases} \quad \pi \otimes \mathrm{d}z - \text{a.s.}$$

By the sample-path left-continuity of γ_ϕ^* in the time variable, (4.4.10) holds for any $t \in \mathbb{R}^+$ with probability 1. Hence, Corollary 4.4.4 implies that $\mathbf{\Gamma}$ admits a left inverse. □

The construction of an invertible mapping is straightforward in a specific case. We recall the notation X^t which stands for the process X stopped at time t for $t \in \mathbb{R}^+$.

Lemma 4.4.6 *Let $\phi \in \mathbb{F}_\infty(\mathcal{P}^{++}(\mathcal{F}^N), \pi)$ be defined by time cases, viz. consider a partition of $\mathbb{R}^+$, $0 = t_0 < t_1 < \ldots < t_k < t_{k+1} = +\infty$ and assume that there exist positive functions $g_j : \mathfrak{M} \times \mathbb{M} \to \mathbb{M}$ such that*

$$\phi(N, s, z) = \sum_{j=0}^{k} \mathbf{1}_{(t_j, t_{j+1}]}(s) g_j(N^{t_j}, z), \quad g_j : \mathfrak{M} \times \mathbb{M} \to \mathbb{M}.$$

Then, $\mathbf{\Gamma}_\phi$ is invertible.

Proof. Let $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$ be defined by time cases. We recall that it is associated to the triplet $(\gamma_\phi, \Gamma_\phi, \phi^\dagger)$ with γ_ϕ π-invertible and defined by time cases. Hence, there exist positive functions $\tilde{g}_j : \mathfrak{M} \times \mathbb{M} \to \mathbb{M}$ such that:

$$\max(\|\tilde{g}_j\|_\infty, \|\frac{1}{\tilde{g}_j}\|_\infty) < +\infty$$

and

$$\phi^\dagger(N, s, z) = \sum_{j=0}^{k} \mathbf{1}_{(t_j, t_{j+1}]}(s) \tilde{g}_j(N^{t_j}, z).$$

Let us construct $\hat{g}_j$ such that for $z \in \mathbb{M}$:

$$\hat{g}_0(N^{t_0}, u) = \frac{1}{g_0(\mathbf{\Gamma}_\phi^{t_0}, z)}.$$

and for $j = 1, \dots, k-1$

$$\hat{g}_j(N, u) = \frac{1}{\tilde{g}_j(\Gamma_\phi^{t_{j-1}}, z)}$$

Then,

$$\psi(N, s, z) := \sum_{j=0}^{k} \mathbf{1}_{(t_j, t_{j+1}]}(s) \hat{g}_j(N, z) = \frac{1}{\phi(\Gamma_\phi, s, \gamma_\phi^*(s, z))}$$

is defined by time cases in $\mathcal{P}^{++}(\mathcal{F}^N)$. It follows by Corollary 4.4.4 that $\mathbf{\Gamma}_\psi \circ \mathbf{\Gamma}_\phi = \mathrm{Id}_{\mathfrak{M}}$ π-a.s.. As in Remark 4.3.18, π_ϕ and π are equivalent which entails, by Theorem 4.3.16, that $\mathbf{\Gamma}_{\phi\#}\pi$ is equivalent to π on $\mathcal{F}_\infty$, and its left invertibility follows. With the same argument as for $\mathbf{\Gamma}_\phi$, $\mathbf{\Gamma}_{\psi\#}$ is equivalent to π. By Lemma 4.4.2, $\mathbf{\Gamma}_\phi \circ \mathbf{\Gamma}_\psi = \mathrm{Id}_{\mathfrak{M}}$ $\mathbf{\Gamma}_{\psi\#}\pi$-a.s., hence π-a.s.. Thus $\mathbf{\Gamma}_\phi$ is also right invertible, which concludes the proof. □

Chapter 5

Applications

As a consequence of the new framework, we devise a new proof of the variational representation of the entropy on the extended Poisson space. Finally, we establish a new criterion for solutions of stochastic differential equations driven by Poisson measures.

5.1 Entropic applications

The left invertibility has important consequences on the relative entropy of $\mathbf{\Gamma}_{\phi\#}\pi$ with respect to the probability measure of reference.

Definition 5.1.1. For μ and μ' two probability measures on $(\mathfrak{M}, \mathcal{F}_\infty)$, the relative entropy of μ' with respect to μ is given by

$$H(\mu' \,|\, \mu) = \begin{cases} \mathbf{E}_{\mu'}\left[\log\left(\left.\frac{\mathrm{d}\mu'}{\mathrm{d}\mu}\right|_{\mathcal{F}_\infty}\right)\right] & \text{if } \mu' \ll \mu \\ +\infty & \text{otherwise.} \end{cases}$$

Consider $\mathfrak{m}$, the smooth, convex, non-negative function defined on $[-1, \infty)$ by

$$\mathfrak{m}(x) = \begin{cases} (x+1)\log(x+1) - x & \text{if } x > -1, \\ 1 & \text{if } x = -1. \end{cases}$$

and $\mathrm{L}_\mathfrak{m}$, the corresponding Orlicz space Adams and Fournier (2003):

$$\mathrm{L}_\mathfrak{m} = \{f \,:\, \mathbb{R}^+ \to [-1, \infty), \int_0^\infty \mathfrak{m}(f(s))\,\mathrm{d}s < \infty\}.$$

Since $\mathfrak{m}(2x) \leq 4\mathfrak{m}(x)$, $\mathrm{L}_\mathfrak{m}$ is a separable Banach space when equipped with the Luxemburg norm, see Adams and Fournier (2003).
In the same vein of (Lassalle, 2012, Proposition 2.1), we obtain the following lemma.

Lemma 5.1.2 *Let $\psi \in \mathcal{P}^{++}(\mathcal{F}^N)$, the following assertions are equivalent:*

(i) There exists $\phi \in \mathbb{F}_2^1(\mathcal{P}^{++}(\mathcal{F}^N), \nu \otimes \mathrm{d}s, \pi_\psi)$ such that $\mathbf{\Gamma}_\phi$ is left invertible.

(ii) $\mathbf{\Gamma}_{\phi\#}\pi = \pi_\psi$ *on* $\mathcal{F}_\infty^N$ *and we have the following identity:*

$$\log\big(\Lambda_\psi \circ \mathbf{\Gamma}_\phi\big) = -\log \Lambda_\phi^\dagger \ \pi - a.s.. \tag{5.1.1}$$

Proof. The left invertibility of $\mathbf{\Gamma}_\phi$ entails that $\mathbf{\Gamma}_\phi \circ \mathbf{\Gamma}_\psi = \mathrm{Id}_{\mathfrak{N}}$ $\mathbf{\Gamma}_{\phi\#}\pi$-a.s. according to Lemma 4.4.2. Recall that the Girsanov theorem says that $\mathbf{\Gamma}_{\psi\#}\pi_\psi = \pi$ hence $\mathbf{\Gamma}_{\phi\#}\pi = \pi_\psi$. According to Lemma 4.4.3, consider Λ_ψ as a process indexed on $\mathbb{R}^+ \times \mathbb{M}$ evaluated at $(t, \mathbb{M})$. For $t \geq 0$, by linearity of $\mathbf{\Gamma}_\phi$ and Lemma 4.4.3,

$$\begin{aligned}
&\log \Lambda_\psi \circ \mathbf{\Gamma}_\phi(N)(t, \mathbb{M}) \\
&\quad = \int_0^t \int \log\Big(\psi(\Gamma_\phi,\, s,\, z)\Big)\, \mathrm{d}\Gamma_\phi(s,z) - \int_0^t \int \Big(\psi(\Gamma,\, s,\, z) - 1\Big)\, \nu(\,\mathrm{d}z)\, \mathrm{d}s \\
&\quad = \int_0^t \int \log\Big(\frac{1}{\phi(N, s, \gamma_\phi^*(N, s, z))}\Big)\, \mathrm{d}\Gamma_\phi(s,z) - \int_0^t \int \Big(\frac{1}{\phi(N, s, \gamma_\phi^*(N, s, z))} - 1\Big)\, \nu(\,\mathrm{d}z)\, \mathrm{d}s, \\
&\quad = \int_0^t \int \log\Big(\phi^\dagger(N, s, z)\Big)\, \mathrm{d}\Gamma_\phi(s,z) - \int_0^t \int \Big(\phi^\dagger(N, s, z) - 1\Big) \nu(\mathrm{d}z)\, \mathrm{d}s.
\end{aligned}$$

using Corollary 4.4.4 and invertibility of γ_ϕ^*. Hence (5.1.1) follows.
For the converse, by (5.1.1), taking the conditional expectation with respect to $\mathcal{F}_t^N$ in both sides of the equality:

$$\begin{aligned}
&\int_0^t \int \Big(\log\Big(\psi(\Gamma,\, s,\, z)\Big) + \log\Big(\phi(N,\, s,\, \gamma_\phi^*(s,z))\Big)\Big)\, \mathrm{d}N(s,z) \\
&\qquad\qquad = \int_0^t \int \Big(\psi(\Gamma,\, s,\, z) + \phi(N, s, \gamma_\phi^*(s,z))\Big)\, \nu(\,\mathrm{d}z)\, \mathrm{d}s
\end{aligned}$$

Equating the jumps yields:

$$\psi(\Gamma,\, s\,,\, z) = \phi(N, s, \gamma_\phi^*(s,z))\ \forall (s,z) \in [0,t] \times \mathbb{M}\ \pi\text{-a.s.}$$

By (4.4.6), it is equivalent to $\mathbf{\Gamma}_\psi \circ \mathbf{\Gamma}_\phi = \mathrm{Id}_{\mathfrak{M}}$ π-a.s..

[Margin comment: $t \to \infty$ at the end and equivalence at once?]

It remains to show that $\mathbf{\Gamma}_{\phi\#}\pi \ll \pi$ along $\mathcal{F}^N$. Eqn (5.1.1) amounts to:

$$\Lambda_\phi^\dagger = \frac{1}{\Lambda_\psi} \circ \mathbf{\Gamma}_\phi.$$

Since $\mathbf{\Gamma}_{\phi\#}\pi = \pi_\psi$, we have:

$$\begin{aligned}
\mathbf{E}_\pi\left[\Lambda_\phi^\dagger\right] &= \mathbf{E}_{\pi_\psi}\left[\frac{1}{\Lambda_\psi}\right] \\
&= \mathbf{E}_\pi\left[\frac{1}{\Lambda_\psi}\Lambda_\psi\right] \quad \text{by the definition of the Radon-Nikodym density } \Lambda_\psi \\
&= 1
\end{aligned}$$

Hence, by lemma 4.3.7, $\mathbf{\Gamma}_{\phi\#}\pi = \pi'_\phi$ is equivalent to π. □

Lemma 5.1.3 *Let $\phi \in \mathbb{F}_2^1(\mathcal{P}(\mathcal{F}^N), \nu \otimes \mathrm{d}s; \pi)$ such that $\mathbf{E}_\pi[\Lambda_\phi] = 1$ and*

$$\mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}\left(\phi^\dagger(s,z) - 1\right)\nu(\mathrm{d}z)\,\mathrm{d}s\right] < \infty.$$

Then,

$$\mathbf{E}_\pi\left[-\log \Lambda^\dagger_{\phi^\dagger}\right] \le \mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}\left(\phi^\dagger(s,z) - 1\right)\nu(\mathrm{d}z)\,\mathrm{d}s\right]. \tag{5.1.2}$$

Proof. From Proposition 4.1.5, as $t \mapsto \Gamma(\cdot, A)$ has compensator

$$t \mapsto \int_0^t \int \phi^\dagger(s,z)\nu(\mathrm{d}z)\,\mathrm{d}s,$$

we have:

$$\begin{aligned}
\mathbf{E}_\pi\left[-\log \Lambda^\dagger_{\phi^\dagger}(t)\right] &= \mathbf{E}_\pi\left[\int_0^t \int \log\big(\phi^\dagger(s,z)\big)\,\mathrm{d}\Gamma_\phi(s,z) - \int_0^t \int (\phi^\dagger(s,z)^{-1} - 1)\nu(\mathrm{d}z)\,\mathrm{d}s\right] \\
&= \mathbf{E}_\pi\left[\int_0^t \int \left(\log(\phi^\dagger(s,z))\phi^\dagger(s,z) - \phi^\dagger(s,z) + 1\right)\nu(\mathrm{d}z)\,\mathrm{d}s\right] \\
&\le \mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}(\phi^\dagger(s,z) - 1)\nu(\mathrm{d}z)\,\mathrm{d}s\right].
\end{aligned}$$

It remains to prove that we can pass to the limit in the left-hand-side. Consider the non-negative, convex function $\psi(x) = x - \log x$. From Fatou's Lemma, we have

$$\begin{aligned}
\mathbf{E}_\pi\left[\psi(\Lambda^\dagger_{\phi^\dagger})\right] &\le \liminf_{t\to\infty} \mathbf{E}_\pi\left[\psi(\Lambda^*_{\phi^\dagger}(t))\right] \\
&\le 1 + \mathbb{E}\left[\int_0^\infty \int \mathfrak{m}(\phi^\dagger(s,z) - 1)\nu(\mathrm{d}z)\,\mathrm{d}s\right].
\end{aligned}$$

This means that the non-negative submartingale $(\psi(\Lambda^\dagger_{\phi^\dagger}(t)),\, t \ge 0)$ is uniformly integrable. Thus,

$$-\log \Lambda^\dagger_{\phi^\dagger}(t) \xrightarrow[t\to\infty]{L^1} -\log \Lambda^\dagger_{\phi^\dagger}.$$

This means that

$$1 + \mathbf{E}_\pi\left[-\log \Lambda^\dagger_{\phi^\dagger}\right] \le 1 + \mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}(\phi^\dagger(s,z) - 1)\nu(\mathrm{d}z)\,\mathrm{d}s\right].$$

The proof is thus complete. □

Theorem 5.1.4 *Let $\phi \in \mathbb{F}_2^1(\mathcal{P}^{++}(\mathcal{F}^N), \nu(\mathrm{d}z) \otimes \mathrm{d}s)$ such that $\mathbf{E}_\pi[\Lambda_\phi] = 1$ and*

$$\mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}\left(\phi^\dagger(s,z) - 1\right)\nu(\mathrm{d}z)\,\mathrm{d}s\right] < \infty.$$

If $\mathbf{\Gamma}_{\phi\#}\pi \ll \pi$, we always have

$$H(\mathbf{\Gamma}_{\phi\#}\pi \,|\, \pi) \le \mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}\left(\phi^\dagger(s,z) - 1\right)\nu(\mathrm{d}z)\,\mathrm{d}s\right]. \tag{5.1.3}$$

Moreover, the map $\mathbf{\Gamma}_\phi$ is left invertible if and only if

$$H(\mathbf{\Gamma}_{\phi\#}\pi \,|\, \pi) = \mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}\left(\phi^\dagger(s,z) - 1\right)\nu(\,\mathrm{d}z)\,\mathrm{d}s\right]. \tag{5.1.4}$$

Proof. Assume that $\mathbf{\Gamma}_{\phi\#}\pi$ is absolutely continuous with respect to π on $\mathcal{F}_\infty$. According to Theorem 4.3.5, there exists $\psi \in \mathbb{F}^1_{2,loc}(\mathcal{P}^{++}(\mathcal{F}^N), \nu \otimes \mathrm{d}s, \mathbf{\Gamma}_{\phi\#}\pi)$ such that

$$\left.\frac{\mathrm{d}\mathbf{\Gamma}_{\phi\#}\pi}{\mathrm{d}\pi}\right|_{\mathcal{F}_t^N} = \Lambda_\psi(N,t).$$

hence for $t \in \mathbb{R}^+$

$$\begin{aligned}\mathbf{E}_\pi\left[f \circ \mathbf{\Gamma}^t_\phi(N)\right] &= \mathbf{E}_\pi\left[f(\Gamma^t)\right] \\ &= \mathbf{E}_\pi\left[f\Lambda_\psi(N,t)\right].\end{aligned}$$

By applying the quasi-invariance Theorem, for $f : \mathfrak{M} \to \mathbb{R}$ bounded and continuous, we have that $f\Lambda_\psi(\cdot,t)$ is bounded and continuous, then

$$\mathbf{E}_\pi\left[f \circ \mathbf{\Gamma}^t_\phi(N)\right] = \mathbf{E}_\pi\left[f \circ \mathbf{\Gamma}^t_\phi(N)\ \Lambda_\psi \circ \mathbf{\Gamma}^t_\phi(N)\ \Lambda^\dagger_{\phi^\dagger}(N,t)\right].$$

Since $|f \circ \mathbf{\Gamma}^t_\phi(N)| \le |f \circ \mathbf{\Gamma}_\phi|$ which is $L^1(\mathfrak{M} \to \mathbb{R}, \pi)$ then, for $t \to \infty$, we obtain

$$\mathbf{E}_\pi\left[f \circ \mathbf{\Gamma}^t_\phi\ \Lambda_\psi \circ \mathbf{\Gamma}_\phi\ \Lambda^\dagger_{\phi^\dagger}\right] \le \mathbf{E}_\pi\left[f \circ \mathbf{\Gamma}_\phi\right].$$

Hence, π-a.s. we have

$$\Lambda_\psi \circ \mathbf{\Gamma}_\phi \times \Lambda^\dagger_\phi \le 1. \tag{5.1.5}$$

It follows that

$$\begin{aligned}0 \le H(\mathbf{\Gamma}_{\phi\#}\pi \,|\, \pi) &= \mathbf{E}_{\mathbf{\Gamma}_{\phi\#}\pi}\left[\log \Lambda_\psi\right] \\ &= \mathbf{E}_\pi\left[\log \Lambda_\psi \circ \mathbf{\Gamma}_\phi\right] \\ &\le -\mathbf{E}_\pi\left[\log \Lambda^\dagger_{\phi^\dagger}\right].\end{aligned} \tag{5.1.6}$$

Then the first part holds.
Assume now that (5.1.4) holds. Then (5.1.6) is an equality as a consequence of (5.1.5), which in turn is an equality. According to lemma 5.1.2, $\mathbf{\Gamma}_\phi$ is left invertible.
Conversely, if $\mathbf{\Gamma}_\phi$ is left invertible. According to the Definition 4.4.1, $\mathbf{\Gamma}_{\phi\#}\pi$ is absolutely continuous with respect to π. Let us denote by $\mathfrak{Z}$ the map such that

$$\mathfrak{Z} \circ \mathbf{\Gamma}_\phi = \mathrm{Id}_{\mathfrak{M}},\ \pi - \text{a.s.}$$

The lemma 5.1.2 entails that:

$$\log \Lambda_\psi \circ \mathbf{\Gamma}_\phi = -\log \Lambda_\phi. \tag{5.1.7}$$

Let

$$R = \frac{\mathrm{d}\mathbf{\Gamma}_{\phi\#}\pi}{\mathrm{d}\pi}.$$

For any $f : \mathfrak{M} \to \mathbb{R}$ continuous and bounded, for any $t > 0$, we have

$$\begin{aligned}\mathbf{E}_\pi[fR] &= \mathbf{E}_\pi[f \circ \mathbf{\Gamma}_\phi] \\ &= \mathbf{E}_\pi\left[(f\Lambda_\psi) \circ \mathbf{\Gamma}_\phi \; \Lambda^\dagger_{\phi^\dagger}\right]\end{aligned}$$

according to (5.1.7) and by the quasi-invariance Theorem.

$$\mathbf{E}_\pi[fR] = \mathbf{E}_\pi[f\Lambda_\psi]$$

It follows that $R = \Lambda_\psi,\ \pi-$a.s. Plug this identity into (5.1.7) to obtain

$$\begin{aligned}H(\mathbf{\Gamma}_{\phi\#}\pi \,|\, \pi) &= \mathbf{E}_\pi[\log R \circ \mathbf{\Gamma}_\phi] \\ &= \mathbf{E}_\pi[\log \Lambda_\psi \circ \mathbf{\Gamma}_\phi] \\ &= \mathbf{E}_\pi\left[-\log \Lambda^\dagger_{\phi^\dagger}\right] \\ &= \mathbf{E}_\pi\left[\int_0^\infty \mathfrak{m}\left(\phi^\dagger(s,z) - 1\right) \nu(\,\mathrm{d}z)\,\mathrm{d}s\right],\end{aligned}$$

using Lemma 5.1.3. The entropic criterion is thus satisfied. □

We reuse the previous notations of Girsanov factor ϕ and associated triplet $(\gamma_\phi, \mathbf{\Gamma}_\phi, \phi^\dagger)$. Let

$$\begin{aligned}\mathcal{P}^{++}_{\mathfrak{m}}(\mathcal{F}^N) &= \left\{\phi \in \mathcal{P}^{++}(\mathcal{F}^N) \text{ and } (\phi^\dagger - 1) \in L^1(\mathfrak{M} \to \mathbb{R}, \pi) \cap \mathrm{L}_{\mathfrak{m}}\right\} \\ \mathcal{P}^{++}_{\infty,\mathrm{pc}}(\mathcal{F}^N) &= \mathcal{P}^{++}_{\infty}(\mathcal{F}^N, \pi) \cap \{\phi \text{ piecewise constant}\} \\ \mathfrak{M}_{\mathfrak{m}}(\mathcal{F}^N) &= \left\{\mu,\ \exists \phi \in \mathcal{P}^{++}_{\mathfrak{m}}(\mathcal{F}^N) \text{ such that } \mu = \mathbf{\Gamma}_{\phi\#}\pi\right\}.\end{aligned}$$

It is well known that there is a Legendre duality between relative entropy and logarithmic Laplace transform Lehec (2013).

Proposition 5.1.5 *Let $f : \mathfrak{M} \to \mathbb{R}$ such that*

$$\mathbf{E}_\pi\left[|f|(1 + e^f)\right] < \infty. \tag{B1}$$

Then,

$$\log \mathbf{E}_\pi\left[e^f\right] = \sup_{\mu \in \mathfrak{M}_{\ll \pi}} \left(\mathbf{E}_\mu[f] - H(\mu \,|\, \pi)\right) \tag{5.1.8}$$

where $\mathfrak{M}_{\ll\pi}$ is the set of probability measures on $\mathfrak{M}$ which are absolutely continuous with respect to π on $\mathcal{F}_\infty$. Furthermore, the supremum is attained at the measure μ_f whose π-density is given by

$$\frac{\mathrm{d}\mu_f}{\mathrm{d}\pi} = \frac{e^f}{\mathbf{E}_\pi[e^f]}. \tag{5.1.9}$$

The theorem of representation of the entropy reads as follows:

Theorem 5.1.6 — Variational representation of the entropy. *Let $f : \mathfrak{M} \to \mathbb{R}$ satisfying* (B1)*. Then,*

$$\log \mathbf{E}_\pi\left[e^f\right] = \sup_{\phi \in \mathcal{P}^{++}_{\mathfrak{m}}(\mathcal{F}^N)} \left(\mathbf{E}_\pi[f(\Gamma_\phi)] - \mathbf{E}_\pi\left[\int_0^\infty \int \mathfrak{m}(\phi^\dagger(s,z) - 1)\nu(\,\mathrm{d}z)\,\mathrm{d}s\right]\right)$$

where $(\gamma_\phi, \mathbf{\Gamma}_\phi, \phi^\dagger)$ is the triplet given by Definition 4.3.8.

Proof. In view of (5.1.8), we evidently have

$$\log \mathbf{E}_\pi \left[e^f\right] \geq \sup_{\mu \in \mathfrak{M}_{\mathfrak{m}}(\mathcal{F}^N)} \left(\mathbf{E}_\mu [f] - H(\mu \,|\, \pi)\right).$$

and

$$\begin{aligned} \sup_{\mu \in \mathfrak{M}_{\ll \pi}} \mathbf{E}_\mu [f] - H(\mu \,|\, \pi) &\geq \sup_{\phi \in \mathcal{P}^{++}_{\mathfrak{m}}(\mathcal{F}^N)} \mathbf{E}_{\mathbf{\Gamma}_{\phi\#\pi}} [f] - H(\mu \,|\, \pi). \\ &\geq \sup_{\phi \in \mathcal{P}^{++}_{\infty,\mathrm{pc}}(\mathcal{F}^N)} \mathbf{E}_{\mathbf{\Gamma}_{\phi\#\pi}} [f] - H(\mathbf{\Gamma}_{\phi\#}\pi \,|\, \pi) \\ &= \sup_{\phi \in \mathcal{P}^{++}_{\infty,\mathrm{pc}}(\mathcal{F}^N)} \mathbf{E}_{\mathbf{\Gamma}_{\phi\#\pi}} [f] - \mathbf{E}_\pi \left[\int_0^\infty \int \mathfrak{m}(\phi^\dagger(s,z) - 1)\nu(\,\mathrm{d}z)\,\mathrm{d}s\right] \end{aligned}$$

by using Theorem 5.1.4 as such $\mathbf{\Gamma}_\phi$ is left invertible in virtue of Lemma 4.4.6. It remains to prove that we can find $(\phi_n,\, n \geq 1)$, a sequence of elements of $\mathcal{P}^{++}_{\infty,\mathrm{pc}}(\mathcal{F}^N, \pi)$ such that

$$\begin{aligned} \mathbf{E}_{\mathbf{\Gamma}_{\phi_n\#\pi}} [f] &\xrightarrow{n\to\infty} \mathbf{E}_{\mu_f} [f] \\ H(\mathbf{\Gamma}_{\phi_n\#}\pi \,|\, \pi) &\xrightarrow{n\to\infty} H(\mu \,|\, \pi) \end{aligned}$$

to conclude. The proof of those limits is analogous to the ones in the proofs of Lemma 2.2. and Theorem 4.11 Zhang (2009-03). □

Remark 5.1.7. Zhang showed a similar result with a different point of view. In his work, the supremum is taken with respect of the Girsanov factor ϕ instead of $\phi^\dagger$. Moreover, f is only bounded. Moreover, we prove that the equality holds if and only if $\mathbf{\Gamma}_\phi$ is left invertible whereas Zhang states a sufficient condition.

5.2 Solutions of SDEs

5.2.1 SDEs driven by Poisson measures

In this section, we specialize the result of right invertibility to characterizing solutions of SDEs under mild assumptions.

Remark 5.2.1. We can in fact start from a "Girsanov factor" to define $\phi^\dagger$ as to define the associated Girsanov transform γ^*_ϕ, and then γ_ϕ as inverse. This point of view will be preferred in that section.

Lemma 5.2.2 *Let $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$ such that:*

$$\phi^\dagger(\omega, s, z) = \frac{\rho((b \circ \alpha)(\omega, s, z))}{\rho(z)} \times ((b \circ \alpha)(\omega, s, \cdot))'(z)$$

with $b : \mathbb{D} \times \mathbb{M} \to \mathbb{R}$ a differentiable and increasing function in z and:

$$\alpha(\omega, s, z) = \left(\sum_{(T_n(\omega), Z_n(\omega)):\, T_n(\omega) < t} Z_n(\omega), z \right).$$

Then,

1. *The process*

$$Y(t) = \int_0^t \int x \, \mathrm{d}\Gamma_\psi(s, x) \tag{5.2.1}$$

is a solution of the SDE

$$Z(t) = \int_0^t \int b(Z(t_-), x) \, \mathrm{d}N(s, x) \tag{5.2.2}$$

if and only if Γ_ψ is the right inverse of Γ_ϕ.

Proof. Let assume $\mathbf{\Gamma}_\phi$ is right invertible with right inverse $\mathbf{\Gamma}_\psi$. Then for $t \in \mathbb{R}^+$,

$$\begin{aligned} \int_0^t \int x \, \mathrm{d}\Gamma_\psi(s, x) &= \int_0^t \int \gamma_\psi(s, x) \, \mathrm{d}N(s, x) \\ &= \int_0^t \int \gamma_\phi^*(\Gamma_\psi, s, x) \, \mathrm{d}N(s, x). \end{aligned}$$

Since

$$(r \circ \gamma_\phi^*(\omega, \cdot))'(s, u) = \phi^\dagger(s, u)\rho(u),$$

we have:

$$\gamma_\phi^*(\omega, s, z) = b \circ \alpha(\omega, s, z).$$

Hence rewriting Eqn (5.2.1), we get:

$$Y(t) = \int_0^t \int b(Y(t_-), x) \, \mathrm{d}N(s, x).$$

Conversely, let assume that there exists a process Y for which (5.2.1) and (5.2.2) hold. Then there exists a predictable process U such that:

$$\int_0^t \int x \, \mathrm{d}\Gamma_\psi(s, x) = \int_0^t \int U(\Gamma_\psi, s, x) \, \mathrm{d}N(s, x).$$

As Γ_ψ is a marked point process, we have: $\gamma_\psi^*(N, s, x) = U(\Gamma_\psi, s, x)\ \pi \otimes \mathrm{d}s \otimes \mathrm{d}x$-a.s.. Hence, it admits a left inverse, i.e. there exists $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$ such that $\gamma_\phi = U$. □

We can generalize this characterization of solutions of SDEs as integrals with respect to a marked point process.

Lemma 5.2.3 *Let $(\gamma_{\phi_m})_{m \in \mathbb{N}}$ a sequence of Girsanov transform of the form:*

$$\gamma_{\phi_m}^*(\omega, s, z) = \sum_{(T_n^m(\omega), Z_n^m(\omega)):\ T_n^m(\omega) < s} \upsilon_m(Z_n^m, z) g(s, Z_n^m).$$

with g a deterministic function and $(\upsilon_m)_{m\in\mathbb{N}}$ a family of differentiable functions with simple limit $(Z_n, z) \mapsto \upsilon(s, Z_n)g(s, Z_n^m)$. , such that $\mathbf{\Gamma}_{\phi_m}$ admits a right inverse $\mathbf{\Gamma}_{\psi_m}$ for $m \in \mathbb{N}$, then:

1. *$(\gamma^*_{\phi_m})_{m\in\mathbb{N}}$ and $(\gamma^*_{\psi_m})_{m\in\mathbb{N}}$ admit limits γ_ϕ and γ_ψ respectively.*

2. *the process Y indexed on $\mathbb{R}^+$:*

$$Y(t) = \int_0^t \int g(s, z)\ \mathrm{d}\Gamma_\psi(s, z).$$

is solution of the SDE:

$$Z(t) = \int_0^t \int g(s, z)$$

Proof. Let $m \in \mathbb{N}$. If $\mathbf{\Gamma}_{\psi_m}$ is the right inverse of $\mathbf{\Gamma}_{\phi_m}$, then $\mathbf{\Gamma}_{\phi_m}$ is the π-left inverse of $\mathbf{\Gamma}_{\psi_m}$ and according to (4.4.7), we have

$$\gamma_{\psi_m}(\omega, s, z) = \gamma^*_{\phi_m}(\Gamma_\psi, s, z) \quad \pi \otimes \mathrm{d}s\ \mathrm{d}z\text{-a.s.}$$

so that for $U \in \mathcal{P}(\mathcal{F}^N)$,

$$\int_0^t \int g(s, z)\ \mathrm{d}\Gamma_\psi(s, z) = \int_0^t \int g(s, \gamma^*_\phi(\Gamma_\psi, s, z))\ \mathrm{d}N(s, z).$$

□

With the same technique, we show another result related to right invertibility.

5.2.2 Markov jump processes

It may be suitable for Markov jump processes and even for the lookalike compensated version, but there are too much involved map compositions to go through with without mentioning the limiting procedure due to the use of indicator which is not differentiable and not necessarily increasing. The huge advantage is that the unicity of SDE solution might be ensured by the abstract notion of invertibility of sample path. Nevertheless, the Girsanov transform should be increasing in z. The curious derivation is that originally the Bismut's approach was designed to tackle non-Markovian processes.

There are 3 ways to characterize a Markov process:

- infinitesimal generator;
- sum of Poisson point processes with a time change;
- use Poisson measure.

We focus on the last interpretation as we can represent almost all Markov processes with a finite number of jump times in this way.

Definition 5.2.4 (Representation of processes in $\mathcal{MP}$).

$$Y^n(t) = Y^n(0) + \int_0^t \int_0^\infty \mathbf{1}_{\{z\le\sigma_0(n-Y^n(s^-))\}} dN_+^n(s,z) - \int_0^t \int_0^\infty \mathbf{1}_{\{z\le\sigma_1 Y^n(s^-))\}} dN_-^n(s,z)$$

where N_+ and N_- are independent Poisson measures. They count for the sign of the increment.

The d-dimensional Markov processes investigated here can be represented as a linear combination of stochastic integrals with respect to a number of Poisson measures. Fix a horizon time $T > 0$ throughout. For any $m \in \mathbb{N}$, any family $(\zeta_1, \dots, \zeta_m)$ of elements of $\mathbb{R}^d$ and any array $(\rho_k)_{1\le k\le m}$ of mappings from $\mathbb{R}^+ \times \mathbb{R}^d$ to $\mathbb{R}$, consider the $\mathbb{R}^d$-valued process X defined as the solution of the SDE

$$Y(t) = Y(0) + \sum_{k=1}^m \left(\int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z\le\rho_k(s,Y(s^-))\}} \,\mathrm{d}N_k(s,z) \right) \cdot \zeta_k, \quad t \le T,$$

where $X(0) \in \mathbb{R}^d$ is fixed, and $(N^k)_{1\le k\le m}$ denote m independent Poisson measures of unit intensity $\mathrm{d}s \otimes \mathrm{d}z$.

Interpretation of right invertibility as another definition of Markov processes, general form (see Decreusefond (1998)). We should highlight in which way it is auto-excitement.

Theorem 5.2.5 *Let $\phi_n \in \mathcal{P}^{++}(\mathcal{F}^N)$ for $n \ge 0$ such that*

$$\gamma_{\phi_n}(\omega, s, z) \xrightarrow{n\to\infty} \gamma_\phi(\omega, s, z) = \varrho(s, \int_0^{s^-} \int \mathbf{1}_{\{u\le z\}} \,\mathrm{d}\omega(s,u))$$

Then, if for $n \ge 1$, $\mathbf{\Gamma}_{\phi_n}$ admits a right inverse $\mathbf{\Gamma}_{\psi_n}$, the following process

$$Y(t) = \int_0^t \int \mathbf{1}_{\{z\le c_0\}} \,\mathrm{d}\Gamma_\psi(s,z)$$

limit of $(Y^n(t))_{n\in\mathbb{N}}$ is a jump Markov process.

Proof. For $g(s,z) = \mathbf{1}_{[0,s]}(s')\mathbf{1}_{(-\infty,c_0]}(z)$, we denote by:

$$Y^n(t) = \int_0^t \int g(s,z) \,\mathrm{d}\mathbf{\Gamma}_{\psi_n}(s,z).$$

Prove we can pass to limit

Then, we have:

$$Y^n(t) = \int_0^t \int \mathbf{1}_{\{z\le\gamma_{\phi_n}(\Gamma_{\psi_n},s,c_0)\}} \,\mathrm{d}N(s,z) \xrightarrow{n\to\infty} \int_0^t \int \mathbf{1}_{\{z\le\varrho(s,Y(s^-))\}} \,\mathrm{d}N(s,z).$$

The limit Y is hence solution of the stochastic differential equation:

$$Y(t) = \int_0^t \int \mathbf{1}_{\{z \leq \varrho(s, Y(s^-))\}} \, \mathrm{d}N(s, z). \tag{5.2.3}$$

□

The Yamada-Watanabe for Poisson measures was proven in Proppe et al. (2014). The Yamada-Watanabe here cannot use the Poisson structure because of the coupling of a random measure and an element of $\mathbb{D}$. In the manuscript, it may be of interest to write a paragraph about the involved techniques. We may need to recall that obtaining strong Markov processes is done through martingales, and more specifically through what is known as martingale problem.

5.3 Strong and weak marked Hawkes processes

5.3.1 Applications

Definition 5.3.1 (Spatial Hawkes process)**.** A family $(Z_t)_{i \in S, t \geq 0}$ of cadlag $(\mathcal{F}_t)_{t \geq 0}$-adapted processes is called a Hawkes process with parameters $(\mathbb{G}, \varphi, \mathbf{h})$ if a.s. for all $i \in S$, all $t \geq 0$

$$Z_t^i = \int_0^t \int_{\mathbb{M}} \mathbb{1}_{\{z \leq h_i(\sum_{j \to i} \int_0^{s^-} \phi_{ji}(s-u) dZ_u^j)\}} \pi^i(dsdz). \tag{5.3.1}$$

5.4 Further works

Can we solve that type of equation?

$$X(t) = \int_0^t u(X_s) \, \mathrm{d}s + W(t) + Z(t), \tag{5.4.1}$$

where W is a Brownian motion and Z is a point process of compensator $y(X, t)$.
What is the criterion for invertibility in that case?
We recall that a Lévy process can be decomposed according to the Lévy-Itô décomposition:

$$X = X^{(1)} + X^{(2)} + X^{(3)}$$

where $X^{(1)}$ is a Brownian motion, $X^{(2)}$ is a compound Poisson process and $X^{(3)}$ is a jump process, square-integrable martingale with almost surely a countable number of jumps on every compact (finite interval). All three processes are independent.
We focus on the case where the components consist only in the non-absolute continuous components for now.
How to handle the jump process?
There are exciting applications of it in the last chapter of Applebaum (2009).
The contribution of a Brownian component on the perturbation γ on the Poisson space is studied in Privault (2003), even without assumption of independence of the Brownian motion and the Poisson measure.

5.5 Future research directions

Bibliography

[1] R. Adamczak, B. Polaczyk, and M. Strzelecki. Modified log-Sobolev inequalities, Beckner inequalities and moment estimates. *Journal of Functional Analysis*, 282(7):109349, 2022.

[2] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003. ISBN 0-12-044143-8.

[3] S. Albeverio and N. Smorodina. A distributional approach to multiple stochastic integrals and transformations of the poisson measure. *Acta Applicandae Mathematica*, 94:1–19, 2006.

[4] A. Anastasiou, A. Barp, F.-X. Briol, B. Ebner, R. E. Gaunt, F. Ghaderinezhad, J. Gorham, A. Gretton, C. Ley, Q. Liu, et al. Stein's method meets computational statistics: A review of some recent developments. *Statistical Science*, 38(1):120–139, 2023.

[5] D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge university press, 2009.

[6] B. Arras, E. Azmoodeh, G. Poly, and Y. Swan. A bound on the wasserstein-2 distance between linear combinations of independent random variables. *Stochastic processes and their Applications*, 129(7):2341–2375, 2019.

[7] B. Arras, J.-C. Breton, A. Deshayes, O. Durieu, and R. Lachièze-Rey. Some recent advances for limit theorems. *ESAIM: Proceedings and Surveys*, 68:73–96, 2020.

[8] R. Arratia, L. Goldstein, L. Gordon, et al. Two moments suffice for Poisson approximations: the Chen-Stein method. *The Annals of Probability*, 17(1):9–25, 1989.

[9] T. Austin. On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probability Surveys*, 5(none):80–145, Jan. 2008. ISSN 1549-5787, 1549-5787. doi: 10.1214/08-PS124. URL `https://projecteuclid.org/journals/probability-surveys/volume-5/issue-none/On-exchangeable-random-variables-and-the-statistics-of-large-graph 10.1214/08-PS124.full`.

[10] E. Azmoodeh, S. Campese, and G. Poly. Fourth moment theorems for Markov diffusion generators. *Journal of Functional analysis*, 266(4):2341–2359, 2014.

[11] E. Azmoodeh, D. Malicet, G. Mijoule, and G. Poly. Generalization of the Nualart–Peccati criterion. *The Annals of Probability*, 44(2):924 – 954, 2016. doi: 10.1214/14-AOP992. URL `https://doi.org/10.1214/14-AOP992`.

[12] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348. Springer Science & Business Media, 2013.

[13] A. D. Barbour. Stein's method and poisson process convergence. *Journal of Applied Probability*, 25(A):175–184, 1988.

[14] A. D. Barbour. Stein's method for diffusion approximations. *Probability theory and related fields*, 84(3):297–322, 1990.

[15] A. D. Barbour, M. Karoński, and A. Ruciński. A central limit theorem for decomposable random variables with applications to random graphs. *Journal of Combinatorial Theory, Series B*, 47(2):125–145, 1989.

[16] A. D. Barbour, L. Holst, and S. Janson. *Poisson approximation*, volume 2. The Clarendon Press Oxford University Press, 1992.

[17] B. B. Bhattacharya, S. Das, and S. Mukherjee. Motif estimation via subgraph sampling: The fourth-moment phenomenon. *The Annals of Statistics*, 50(2):987–1011, 2022.

[18] K. Bichteler and J. Jacod. Calcul de Malliavin pour les diffusions avec sauts : existence d'une densité dans le cas unidimensionnel. *Séminaire de probabilités de Strasbourg*, 17: 132–157, 1983. URL `https://eudml.org/doc/113430`.

[19] K. Bichteler, J. Gravereaux, and J. Jacod. *Malliavin Calculus for Processes with Jumps*. Stochastic Monographs : Theory and Applications of Stochastic Processes, Vol 2. Gordon and Breach Science Publishers, 1987. ISBN 9782881241857. URL `https://books.google.fr/books?id=5AvvAAAAMAAJ`.

[20] J. Bismut. Calcul des variations stochastique et processus de sauts. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 63(2):147–235, 1983. ISSN 0044-3719, 1432-2064. doi: 10.1007/BF00538963. URL `http://link.springer.com/10.1007/BF00538963`.

[21] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca. Network analysis in the social sciences. *science*, 323(5916):892–895, 2009.

[22] A. V. Bulinski. Conditional central limit theorem. *Theory of Probability & Its Applications*, 61(4):613–631, 2017.

[23] S. Chatterjee et al. A new method of normal approximation. *The Annals of Probability*, 36(4):1584–1610, 2008.

[24] L. H. Chen. Poisson approximation for dependent trials. *The Annals of Probability*, pages 534–545, 1975.

[25] L. H. Chen. Stein's method of normal approximation: Some recollections and reflections. *The Annals of Statistics*, 49(4):1850–1863, 2021.

[26] L. H. Chen, Q.-M. Shao, et al. Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli*, 13(2):581–599, 2007.

[27] L. H. Chen, L. Goldstein, and Q.-M. Shao. *Normal approximation by Stein's method*. Springer Science & Business Media, 2010.

[28] K. L. Chung. *A course in probability theory*. Academic press, 1974.

[29] M. Coulson, R. E. Gaunt, and G. Reinert. Poisson approximation of subgraph counts in stochastic block models and a graphon model. *ESAIM: Probability and Statistics*, 20: 131–142, 2016.

[30] L. Coutin and L. Decreusefond. Invertibility of functionals of the Poisson process and applications. In preparation, 2023.

[31] P. De Jong. A central limit theorem for generalized multilinear forms. *Journal of Multivariate Analysis*, 34(2):275–289, 1990.

[32] P. De Jong. A central limit theorem with applications to random hypergraphs. *Random Structures & Algorithms*, 8(2):105–120, 1996.

[33] P. a. de Jong. *Central Limit Theorems for Generalized Multilinear Forms*, volume 61 of *CWI tract*. Centrum voor Wiskunde en Informatica (Amsterdam, Netherlands), 1989.

[34] L. Decreusefond. Perturbation Analysis and Malliavin Calculus. *Annals of Applied Probability*, 8(2):496–523, 1998.

[35] L. Decreusefond. The Stein-Dirichlet-Malliavin method. *ESAIM: Proceedings and Surveys*, 51:49–59, 2015.

[36] L. Decreusefond. *Selected Topics in Malliavin Calculus: Chaos, Divergence and So Much More*. Springer Nature, 2022.

[37] L. Decreusefond and H. Halconruy. Malliavin and Dirichlet structures for independent random variables. *Stochastic Processes and their Applications*, 129(8):2611–2653, Aug 2019. doi: 10.1016/j.spa.2018.07.019. URL `https://hal.archives-_ouvertes.fr/hal-_01565240`.

[38] J.-F. Delmas, J.-S. Dhersin, and M. Sciauveau. Asymptotic for the cumulative distribution function of the degrees and homomorphism densities for random graphs sampled from a graphon. *Random Structures & Algorithms*, 58(1):94–149, 2021.

[39] P. S. Dey and G. Terlov. Stein's method for conditional central limit theorem. *The Annals of Probability*, 51(2):723–773, 2023.

[40] C. Döbler. Normal approximation via non-linear exchangeable pairs. *Latin American Journal of Probability and Mathematical Statistics*, 20:167, 01 2023. doi: 10.30757/ALEA.v20-|08.

[41] C. Döbler, G. Peccati, et al. Quantitative de Jong theorems in any dimension. *Electronic Journal of Probability*, 22, 2017.

[42] C. Döbler, G. Peccati, et al. The fourth moment theorem on the Poisson space. *Annals of Probability*, 46(4):1878–1916, 2018.

[43] C. Döbler, K. Krokowski, et al. On the fourth moment condition for Rademacher chaos. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 55, pages 61–97. Institut Henri Poincaré, 2019.

[44] C. Döbler, G. Peccati, et al. Quantitative CLTs for symmetric U-statistics using contractions. *Electronic Journal of Probability*, 24, 2019.

[45] C. Döbler, M. J. Kasprzak, and G. Peccati. Functional convergence of sequential U-processes with size-sependent kernels. *The Annals of Applied Probability*, 32(1):551–601, 2022.

[46] R. M. Dudley. *Real Analysis and Probability*. Cambridge studies in advanced mathematics. Cambridge University Press, 2002. ISBN 9780521809726. URL `https://books.google.fr/books?id=J2XHngEACAAJ`.

[47] M. Duerinckx. On the size of chaos via Glauber calculus in the classical mean-field dynamics. *Communications in Mathematical Physics*, pages 1–41, 2021.

[48] N. T. Dung. Poisson and normal approximations for the measurable functions of independent random variables. *arXiv preprint arXiv:1807.10925*, 2018.

[49] N. T. Dung. Explicit rates of convergence in the multivariate CLT for nonlinear statistics. *Acta Mathematica Hungarica*, 158(1):173–201, 2019.

[50] N. T. Dung. Rates of convergence in the central limit theorem for nonlinear statistics under relaxed moment conditions. *Acta Mathematica Vietnamica*, pages 1–26, 2021.

[51] B. Efron and C. Stein. The jackknife estimate of variance. *The Annals of Statistics*, pages 586–596, 1981.

[52] G. Elek and B. Szegedy. A measure-theoretic approach to the theory of dense hypergraphs. *Advances in Mathematics*, 231(3-4):1731–1772, 2012.

[53] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc, 1986.

[54] X. Fang and A. Röllin. Rates of convergence for multivariate normal approximation with applications to dense graphs and doubly indexed permutation statistics. *Bernoulli*, 21(4): 2157–2189, 2015.

[55] M. Fujisaki and T. Komatsu. Generalized Girsanov Transform of Processes and Zakai Equation with Jumps. *Journal of Stochastic Analysis*, 2(3):17, Sept. 2021. ISSN 2689-6931. doi: 10.31390/josa.2.3.17. URL `https://digitalcommons.lsu.edu/josa/vol2/iss3/17`.

[56] L. Goldstein et al. Bounds on the constant in the mean central limit theorem. *Annals of Probability*, 38(4):1672–1689, 2010.

[57] F. Gotze. On the rate of convergence in the multivariate clt. *The Annals of Probability*, pages 724–739, 1991.

[58] W. Grzenda and W. Zieba. Conditional central limit theorem. In *Int. Math. Forum*, volume 3, pages 1521–1528, 2008.

[59] W. Hoeffding. A Class of Statistics with Asymptotically Normal Distribution. *The Annals of Mathematical Statistics*, pages 293–325, 1948.

[60] C. Houdré, N. Privault, et al. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, 8(6):697–720, 2002.

[61] J. Jacod. Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31(3):235–253, Sept. 1975. ISSN 1432-2064. doi: 10.1007/BF00536010. URL `https://doi.org/10.1007/BF00536010`.

[62] J. Jacod. *Calcul Stochastique et Problèmes de Martingales*. Springer-Verlag, 1979.

[63] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer-Verlag Berlin Heidelberg GmbH, 2 edition, 2003.

[64] S. Janson and K. Nowicki. The asymptotic distributions of generalized U-statistics with applications to random graphs. *Probability theory and related fields*, 90(3):341–375, 1991.

[65] S. Janson, T. Luczak, and A. Rucinski. Random graphs, 2000.

[66] O. Kallenberg. *Foundations of Modern Probability*. Springer, 1997.

[67] O. Kallenberg. *Foundations of Modern Probability*. Springer Science & Business Media, 2002.

[68] G. Kaur and A. Röllin. Higher-order fluctuations in dense random graph models. *Electronic Journal of Probability*, 26:1–36, 2021.

[69] D. Knoke and S. Yang. *Social network analysis*, volume 154. Sage Publications, 2019.

[70] K. Krokowski, A. Reichenbachs, C. Thäle, et al. Discrete Malliavin–Stein method: Berry–Esseen bounds for random graphs and percolation. *The Annals of Probability*, 45(2): 1071–1109, 2017.

[71] R. Lachièze-Rey, G. Peccati, et al. New Berry–Esseen bounds for functionals of binomial point processes. *The Annals of Applied Probability*, 27(4):1992–2031, 2017.

[72] R. Lassalle. *Stochastic Invertibility and Related Topics*. PhD thesis, Telecom ParisTech, 2012.

[73] M. Ledoux et al. Chaos of a Markov operator and the fourth moment condition. *Annals of Probability*, 40(6):2439–2459, 2012.

[74] J. Lehec. Representation formula for the entropy and functional inequalities. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 49(3):885–899, 2013. ISSN 0246-0203. doi: 10.1214/11-|aihp464.

[75] L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[76] C. McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1): 148–188, 1989.

[77] M. E. Newman, A.-L. E. Barabási, and D. J. Watts. *The structure and dynamics of networks*. Princeton university press, 2006.

[78] I. Nourdin and G. Peccati. Stein's method on Wiener chaos. *Probability Theory and Related Fields*, 145(1-2):75–118, 2009.

[79] I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus: from Stein's method to universality.* Number 192 in Cambridge Tracts in Mathematics. Cambridge University Press, 2012.

[80] I. Nourdin, G. Peccati, and G. Reinert. Second order Poincaré inequalities and CLTs on Wiener space. *Journal of Functional Analysis*, 257(2):593–609, 2009.

[81] K. Nowicki and J. C. Wierman. Subgraph counts in random graphs using incomplete U-statistics methods. *Discrete Mathematics*, 72(1-3):299–310, 1988.

[82] D. Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer, 2006.

[83] D. Nualart, G. Peccati, et al. Central limit theorems for sequences of multiple stochastic integrals. *Annals of Probability*, 33(1):177–193, 2005.

[84] G. Peccati. Hoeffding-ANOVA decompositions for symmetric statistics of exchangeable observations. *The Annals of Probability*, 32(3):1796–1829, 2004.

[85] N. Privault. Girsanov theorem for anticipative shifts on Poisson space. *Probability Theory and Related Fields*, 104(1):61–76, Mar. 1996. ISSN 1432-2064. doi: 10.1007/BF01303803. URL `https://doi.org/10.1007/BF01303803`.

[86] N. Privault. Quasi-invariance for Lévy processes under anticipating shifts. In *Stochastic Analysis and Related Topics VIII: Silivri Workshop in Gazimagusa (North Cyprus), September 2000*, pages 181–202. Springer, 2003.

[87] N. Privault and G. Serafin. Normal approximation for sums of discrete U-statistics - Application to Kolmogorov bounds in random subgraph counting. *arXiv preprint arXiv:1806.05339*, 2018.

[88] N. Privault and G. Serafin. Berry-Esseen bounds for functionals of independent random variables. *Electronic Journal of Probability*, 27:1–37, 2022.

[89] N. Privault, G. Serafin, et al. Normal approximation for sums of discrete u-statistics–application to Kolmogorov bounds in random subgraph counting. *Bernoulli*, 26(1):587–615, 2020.

[90] N. Privault et al. Stochastic analysis of Bernoulli processes. *Probability Surveys*, 5:435–483, 2008.

[91] C. Proppe, J. Villarroel, and H. Zhao. Yamada-watanabe theorem for stochastic evolution equation driven by poisson random measure. *ISRN Probability and Statistics*, 2014:982190, 2014. doi: 10.1155/2014/982190. URL `https://doi.org/10.1155/2014/982190`.

[92] S. Rachev. *Probability Metrics and the Stability of Stochastic Models.* Wiley Series in Probability and Statistics - Applied Probability and Statistics Section. Wiley, 1991. ISBN 9780471928775. URL `https://books.google.fr/books?id=5grvAAAAMAAJ`.

[93] M. Raginsky and I. Sason. *Concentration of Measure Inequalities in Information Theory, Communications, and Coding.* Now Foundations and Trends, 2014.

[94] B. P. Rao. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61(2):441–460, 2009.

[95] G. Reinert. Couplings for normal approximations with Stein's method. *DIMACS Ser. Discrete Math. Theoret. Comput. Sci*, 41:193–207, 1998.

[96] G. Reinert, I. Nourdin, G. Peccati, et al. Stein's method and stochastic analysis of Rademacher functionals. *Electronic Journal of Probability*, 15:1703–1742, 2010.

[97] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields*, 78(1):1–10, 1988.

[98] A. Röllin. Kolmogorov bounds for the normal approximation of the number of triangles in the Erdos-Rényi random graph. *Probability in the Engineering and Informational Sciences*, 36(3):747–773, 2022. doi: 10.1017/S0269964821000061.

[99] H. Sambale and A. Sinulis. Logarithmic Sobolev inequalities for finite spin systems and applications. *Bernoulli*, 26(3):1863–1890, 2020.

[100] J. Schuppen and E. Wong. Transformation of Local Martingales Under a Change of Law. *The Annals of Probability*, 2(5):879–888, Oct. 1974. ISSN 0091-1798, 2168-894X. doi: 10.1214/aop/1176996554. URL `https://projecteuclid.org/journals/annals-_of-_probability/volume-_2/issue-_5/Transformation-_of-_Local-_Martingales-_Under-_a-_Change-_of-_Law/10.1214/aop/1176996554.full`.

[101] A. Skorokhod. On the differentiability of measures which correspond to stochastic processes. i. processes with independent increments. *Theory of Probability & Its Applications*, 2(4):407–432, 1957.

[102] C. Stein et al. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*. The Regents of the University of California, 1972.

[103] T. Temčinas, V. Nanda, and G. Reinert. Multivariate Central Limit Theorems for Random Clique Complexes, June 2022. URL `http://arxiv.org/abs/2112.08922`. arXiv:2112.08922 [math] type: article.

[104] I. Tyurin. New estimates of the convergence rate in the Lyapunov theorem, 2009.

[105] A. S. Üstünel. Entropy, invertibility and variational calculus of adapted shifts on Wiener space. *Journal of Functional Analysis*, 257(11):3655–3689, Dec. 2009. ISSN 0022-1236. doi: 10.1016/j.jfa.2009.03.015.

[106] A. S. Üstünel. Variational calculation of Laplace transforms via entropy on Wiener space and applications. *Journal of Functional Analysis*, 267(8):3058 – 3083, 2014. ISSN 0022-1236. doi: https://doi.org/10.1016/j.jfa.2014.07.006. URL `http://www.sciencedirect.com/science/article/pii/S0022123614002730`.

[107] K. Yosida. *Functional Analysis*. Classics in Mathematics. Springer Berlin Heidelberg, 1995. ISBN 9783540586548.

[108] D.-M. Yuan, L.-R. Wei, and L. Lei. Conditional central limit theorems for a sequence of conditional independent random variables. *Journal of the Korean Mathematical Society*, 51(1):1–15, 2014.

[109] X. Zhang. Clark-Ocone formula and variational representation for Poisson functionals. *The Annals of Probability*, 37(2):506–529, 2009-03. doi: 10.1214/08-aop411.

[110] Z.-S. Zhang. Berry–Esseen bounds for generalized U-statistics. *Electronic Journal of Probability*, 27:1–36, 2022.

[111] G. Zheng. A Peccati-Tudor type theorem for Rademacher chaoses. *ESAIM: Probability and Statistics*, 23:874–892, 2019.

[112] A. Üstünel and M. Zakai. *Transformation of Measure on Wiener Space*. Springer Science & Business Media, 01 2000. ISBN 978-3-642-08572-7. doi: 10.1007/978-3-662-13225-8.

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : Contributions à l'analyse stochastique pour structures sans propriété de diffusion

Mots clés : Calcul de Malliavin, Méthode de Stein, Variables aléatoires conditionnellement indépendantes, Mesure de Poisson, Théorème de Girsanov, Critère entropique, Equations différentielles stochastiques

Résumé : Cette thèse a pour sujet l'étude de structures sans propriété de diffusion. Nous nous intéressons à deux classes de telles structures.
Le premier sujet traite du calcul de Malliavin pour les variables aléatoires conditionnellement indépendantes qui est un cas de calcul de Malliavin discret. Il généralise aussi celui théorisé sur des produits dénombrables d'espaces de probabilité, pour les variables aléatoires indépendantes. Dans notre cas, l'intérêt d'un tel calcul est de venir compléter des résultats d'analyse stochastique avec des preuves d'inégalités fonctionnelles (inégalité de Poincaré, inégalité de McDiarmid) et de théorèmes limites. Une des applications phares est la détermination de la vitesse de convergence de théorèmes centraux limites via la méthode de Stein. En combinant le calcul de Malliavin avec la structure de Dirichlet sous-jacente aux variables aléatoires, nous obtenons une formule d'intégration par parties cruciale pour déterminer des bornes supérieures sur les vitesses de convergence. Nous montrons des théorèmes limites quantitatifs, dont un théorème de quatrième moment avec reste. En particulier, nous discutons d'une application à la normalité asymptotique du comptage de motifs dans des hypergraphes aléatoires échangeables.
Le deuxième sujet étudie les fonctionnelles d'une mesure de Poisson en utilisant la notion d'inversibilité de transformations de cette mesure sur l'espace échantillon des mesures aléatoires. Nous utilisons l'identification de ces mesures et des processus ponctuels marqués associés. Les transformations inversibles sont obtenues via le théorème de Girsanov, en respectant l'absolue continuité par rapport à la mesure de référence. Il en résulte un critère entropique pour l'inversibilité des transformations. Enfin, nous faisons le lien avec les équations différentielles stochastiques dirigées par des mesures de Poisson.

Title : Contributions to stochastic analysis for non-diffusive structures

Keywords : Malliavin calculus, Stein's method, Conditionally independent random variables, Poisson measure, Girsanov's theorem, Entropic criterion, Stochastic differential equations

Abstract : This thesis is concerned with the study of non-diffusive structures. We focus on two classes of such structures.
The first subject deals with Malliavin calculus for conditionally independent random variables, which is a special case of discrete Malliavin calculus. It also generalizes the calculus that has been developed for countable products of probability spaces, for independent random variables. In our case, the interest of such a calculus is to complement results in stochastic analysis with proofs of functional inequalities (Poincaré inequality, McDiarmid's inequality) and limit theorems. One of the main applications is the determination of the convergence rate of central limit theorems via the Stein method. By combining Malliavin calculus with the underlying Dirichlet structure of the random variables, we obtain an integration by parts formula which is key to the derivations of so-called Stein bounds of the rates of convergence. We show quantitative limit theorems, including a fourth moment theorem with remainder. In particular, we discuss an application to the asymptotic normality of motif counting in exchangeable random hypergraphs.
The second subject studies functionals of a Poisson measure using the notion of invertibility of transformations of that measure on the sample space of random measures. We use the identification of these measures and the associated marked point processes. Invertible transformations are obtained via the Girsanov's theorem, respecting absolute continuity with respect to the reference measure. This results in an entropy criterion for the invertibility of transformations. Finally, we make the connection with stochastic differential equations driven by Poisson measures.

Institut Polytechnique de Paris
91120 Palaiseau, France